Contrôle stochastique et applications à la couverture d'options en présence d'illiquidité: Aspects théoriques et numériques

Benjamin Bruder

Université Paris VII - Denis Diderot
UFR de Mathématiques
Année 2008

Thèse

pour obtenir le titre de
Docteur de l'Université Paris 7
Spécialité : Mathématiques Appliquées

présentée par
Benjamin BRUDER

Contrôle stochastique et applications à la couverture d'options en présence d'illiquidité : Aspects théoriques et numériques

---

Directeur de thèse
Pr. PHAM Huyên

Soutenue publiquement le 17 Janvier 2008, devant le jury composé de :

CONT Rama, Professeur, Ecole Polytechnique et Columbia University
EL KAROUI Nicole, Professeur, Ecole Polytechnique
JAMET Guillaume, Société Générale Asset Manangement
PHAM Huyên, Professeur, Université Paris 7
TALAY Denis, Directeur de Recherche, INRIA
TOUZI Nizar, Professeur, Ecole Polytechnique

au vu des rapports de :

CONT Rama, Professeur, Ecole Polytechnique et Columbia University
SONER Mete, Professeur, London School of Economics

Je tiens tout d'abord à remercier mon directeur de thèse, Huyên Pham. Sans lui, cette thèse n'aurait pu ni commencer ni aboutir. Je le remercie de ses nombreux conseils, de son attention, ainsi que de sa très grande disponibilité. Il m'a permis d'aborder les problèmes mathématiques avec la rigueur indispensable.
Je tiens aussi à remercier aussi Guillaume Jamet pour m'avoir encadré à la SGAM, notamment pour les problématiques intéressantes qu'il m'a fournies et qui son à l'origine du travail accompli ici.
Je suis très reconnaissant envers Rama Cont et Mete Soner pour avoir accepté de rédiger les rapports de cette thèse, ainsi que Nicole El Karoui, Denis Talay et Nizar Touzi pour avoir accepté de faire partie du jury.
Je remercie chaleureusement Olivier Bokanowski, Stefania Maroso et Hasnaa Zidani pour leur aide et leur collaboration dans le domaine numérique. Ils ont su me faire apprécier cette discipline.
Je souhaite aussi remercier Nicolas Gaussel qui est à l'initiative de cette thèse.
J'aimerais remercier le personnel de la SGAM, notamment Marc Barton Smith, Vivien Brunel, Rabih Chaar, Paul Demey, Philippe Dumont, Walid Gueriri, Guillaume Lasserre, Richard Rouge et Francois Soupe. J'ai pu profiter de leur compétence et de leur sympathie pendant ces trois années.
Je remercie enfin mes amis thésards de Chevaleret pour leur amitié, en particulier Katia Mezziani, Vathana Ly Vath, Tu Nguyen, Francois Simenhaus, Vincent Vargas et Thomas Willer.

Table des matières

Introduction générale

Introduction

Durant les années 1970, les mathématiques financières ont connu une révolution grâce à l'utilisation du calcul stochastique. D'une part par l'introduction de la notion de couverture d'option par Black et Scholes [16], et d'autre part par l'article de Merton [53] appliquant les techniques du contrôle stochastique au problème d'investissement optimal en temps continu. Depuis, de nombreux travaux ont cherché à enrichir ces problèmes en tenant compte d'éléments tels que l'incomplétude des marchés, les coûts de transaction ou l'illiquidité de certains actifs. D'autres recherches ont eu pour but d'affiner la dynamique des prix des actifs financiers, afin de les rendre plus cohérents avec l'observation des données de marchés.

Le marché des produits dérivés s'est considérablement développé depuis cette époque, particulièrement pour les grands indices boursiers tels que le S&P500 et l'Eurostoxx 50. Certaines options Européennes, telles que les calls et des puts sur ces indices, sont devenus des actifs pour lesquels il existe un marché organisé. Pour ces grands indices, les banques cherchent à fournir des payoffs de plus en plus complexes. Les options Européennes sont alors utilisées par les banques pour couvrir les options exotiques contre le risque de variation de volatilité de l'indice. Il s'est récemment développé une nouvelle classe de produits liquides : les swaps de variance (le VIX par exemple). Ils consistent à échanger un flux fixe contre la variance réalisée de l'indice considéré. Ceux ci pourraient supplanter les options Européennes en tant que référence de produit de couverture du risque de volatilité. Cependant, il existe encore de nombreux actifs pour lesquels il n'existe pas de marché d'options liquides. Par exemple, sur des indices boursiers plus petits, les options sont sujettes à de forts coûts de transaction. A l'extrême, les options sur fonds d'investissements tels que les mutual funds et les hedge funds sont vendues de gré à gré sans qu'il n'existe de cotations ou de de liquidité pour de telles options. Nous sommes donc, dans ces cas là, dans le cadre de marchés imparfaits.
La théorie du contrôle stochastique est un excellent outil pour valoriser les options en marché imparfait. En effet, elle permet de trouver les stratégies d'investissement permettant la minimisation d'un critère de risque donné. Ceci est généralement fait en caractérisant l'espérance du risque minimal comme solution d'une équation aux dérivées partielles de type Hamilton Jacobi Bellman. Pour une introduction générale à cette discipline, on pourra se référer au livre de Soner et Fleming [38] ou de Pham [59] dans le cas des applications à la finance. Ensuite, il reste à résoudre numériquement cette EDP pour obtenir de bonnes solutions approchées. Dans cette thèse, nous utiliserons ces techniques afin obtenir des résultats de pricing et de couverture d'options dans le cas de marchés imparfaits, dans l'optique de pouvoir les appliquer aux options sur mutual funds et hedge funds. Les aspects à étudier pour la couverture de telles options sont nombreux. Cette thèse est constituée de trois parties indépendantes. La première concerne la surréplication sous contraintes gamma. La deuxième porte sur une approche du risque de volatilité. Enfin la troisième et dernière partie étudie les problèmes de contrôle avec retard.

En ce qui concerne la liquidité des instruments de couverture, on peut tout d'abord s'intéresser aux coûts de transaction. Ceux ci rendent le marché incomplet et, mis a part quelques cas particuliers, la réplication parfaite des produits dérivés est impossible. Dans ce cas, on ne peut plus définir un prix unique par absence d'opportunité d'arbitrage. La

situation est telle qu'il existe un intervalle de prix non arbitrables. Les travaux de Cvitanic, Pham et Touzi [28] ainsi que ceux de Ben Tahar et Bouchard [12] étudient ces bornes d'arbitrage lorsque l'actif sous jacent d'une option Européenne est soumis à des coûts de transaction. Ils étudient pour cela le problème de surréplication de l'option, développé par El Karoui et Quenez [36], en y incluant les coûts de transaction. Une autre possibilité est d'étudier le prix par indifférence de l'option, introduit par Hodges [45]. Celui ci donne le prix pour lequel un agent caractérisé par une fonction d'utilité donnée sera indifférent à vendre ou non une certaine quantité d'options. Dans leur article, Davis Panas et Zariphopoulou [31] utilisent un problème de contrôle singulier afin de pouvoir caractériser par une inéquation variationnelle l'utilité optimale de l'agent ayant vendu l'option. Ceci permet alors d'obtenir ce prix d'indifférence.
La première partie de cette thèse concerne une approche indirecte aux coûts de transaction via la surréplication avec contraintes de gamma, développée par Cheridito, Soner et Touzi [63], [23], [65]. Le problème consiste à prendre comme point de départ un marché complet, mais à réduire l'ensemble des stratégies de couverture admissibles. L'idée est de décomposer la dynamique instantanée de la quantité d'actif détenue dans le portefeuille de l'agent en une partie Brownienne et une partie à variation finie. La contrainte gamma consiste alors à contraindre l'intégrand de la partie Brownienne à rester dans un certain ensemble convexe. En d'autre termes, c'est une contrainte sur la volatilité de la quantité d'actif sous jacent détenue par l'agent. Or, la partie Brownienne ayant presque sûrement une variation infinie, elle implique théoriquement des coûts de transaction infinis. Intuitivement, la contrainte gamma est donc susceptible de réduire en pratique les coûts de transaction liés à la couverture d'option. Dans cette thèse, nous étudierons plus particulièrement le cas particulier dégénéré pour lequel on annule la partie Brownienne. On contraint donc les stratégies à être à variation finie. Ceci impliquera ainsi un montant fini de coût de transactions.

Dans une deuxième partie, nous nous intéressons au paramètre de volatilité. Celui ci a lui aussi été l'objet d'attention depuis le krach boursier de 1987. En effet, c'est alors qu'est apparu de manière significative le smile de volatilité. Or, celui ci entre en contradiction avec l'hypothèse de diffusion Brownienne avec volatilité constante faite par Black et Scholes. Plusieurs types de modèles ont cherché à expliquer ce comportent, notamment les modèles à volatilité stochastique de Heston [43] ou de Hull et White [47]. Cependant, pour espérer couvrir parfaitement une option dans un tel cadre de travail, il faut disposer d'un instrument de couverture de la volatilité parfaitement liquide mais aussi pouvoir supposer les paramètres du modèle fixés et parfaitement connus. Or, en pratique, ces conditions ne sont pas toujours vérifiées.
Lorsqu'il n'existe aucun instrument de couverture de la volatilité, dans le cadre d'options sur fonds par exemple, plusieurs approches sont possibles. Tout d'abord les méthodes standard de marchés incomplets. On peut utiliser le pricing par indifférence en résolvant des problèmes de contrôle réguliers dans le cadre markovien des équation HJB, comme dans l'article de Musiela et Zariphopoulou [54]. Plusieurs travaux, par exemple ceux de Rouge et El Karoui [61], ont aussi cherché à trouver le prix par indifférence dans le cadre des équations différentielles stochastiques rétrogrades. On peut enfin, chercher la stratégie de couverture donnant la variance minimale pour le prix de couverture. Pour une vue générale du sujet, on pourra se référer aux articles de Pham [57] et de Schweizer [62]. Cependant, dans tous ces cas, il est nécessaire de supposer la dynamique de la volatilité parfaitement connue. Dans le cadre général d'incertitude sur le modèle, on pourra par exemple se référer

aux travaux de Cont [25]. Dans le cadre de l'incertitude sur le paramètre de volatilité, en revanche, il est possible d'utiliser les techniques de surréplication. Les travaux de Cvitanic, Pham et Touzi [29] ont trouvé le prix de surréplication dans le cadre d'une volatilité susceptible d'évoluer dans l'intervalle des nombres réels positifs. Cependant, le prix obtenu est alors un prix d'arbitrage statique trivial qui est inutilisable en pratique. Avellaneda, Levy et Paras [3], puis Gozzi et Vargoliu [40] ont quand à eux considéré le cas où la volatilité est supposée rester dans un intervalle donné. Nous chercherons dans la troisième partie de cette thèse à généraliser ces méthodes, en considérant le prix de surréplication pour une volatilité non bornée, mais nous chercherons à éviter les prix d'arbitrage triviaux en introduisant des pertes tolérées en fonction de la trajectoire de volatilité.

Enfin, dans la troisième partie, nous considérerons un type d'imperfection peu étudié : le retard à l'exécution des ordres. En effets les ordres d'achat et de vente de parts de hedge funds doivent être déclarés un ou plusieurs mois avant de pouvoir être exécutés. Ceci entraîne évidemment des complications lorsqu'il s'agit de couvrir des options sur un tel sous jacent. Ce problème peut être vu de deux manières différentes. Follmer et Schweizer [39] considèrent la couverture donnant une erreur de variance minimale dans le cas d'un retard d'information. La solution est alors de projeter au sens de L^2 la stratégie optimale en information parfaite sur la filtration retardée. Bar-Illan et Sulem [4], quand à eux, résolvent un problème de contrôle stochastique avec retard à l'exécution, en horizon infini avec une dynamique linéaire. Récemment, Oksendal et Sulem [55] ont montré l'équivalence des problème de retard d'information et d'exécution. Cela aboutit à la résolution du problème lorsque la dynamique des processus contrôlés satisfait une certaine hypothèse, restrictive pour certaines applications en finance. Nous chercherons dans une dernière partie à étudier les problèmes de contrôles avec retard permettant de résoudre notre problème de couverture d'option avec retard.

1 Présentation des résultats de la première partie

1.1 Chapitre 1 : Problème de contraintes gamma et obtention de l'équation caractéristique

On considère un espace probabilisé $(\Omega, \mathcal{F}, P)$. On considère un marché à trois actifs, et un horizon T. Le premier actif est l'actif sans risque. Comme nous considérerons les prix actualisés, on suppose que son prix est constant et égal à 1. Le deuxième actif, dont le prix sera noté s, représente par exemple le prix d'un indice boursier. On considère qu'il n'y a pas de dividendes. Enfin, le troisième actif représente, par exemple, le prix d'un contrat swap de variance ayant pour sous jacent le deuxième actif. Son prix est noté x. Sa maturité est supérieure à T pour qu'il n'y aie pas de dégénérescence de son prix en T. On suppose que ces deux actifs suivent la dynamique :

$$\left\{\begin{array}{c} S_t^{t,s,x} = s \ \ , \ \ X_t^{t,s,x} = x \\ dS_u^{t,s,x} = S_u^{t,s,x}\sigma\left(t, X_u^{t,s,x}\right) dW_u^1 \\ dX_u^{t,s,x} = -\mu\left(u, X_u^{t,s,x}\right) du + \zeta\left(t, X_u^{t,s,x}\right) dW_u^1 + \xi\left(t, X_u^{t,s,x}\right) dW_u^2 \end{array}\right.$$

Où (W^1, W^2) est un mouvement Brownien standard de dimension 2. L'actif x distribue continûment des dividendes $\mu(t,x)$. L'hypothèse importante dans ce modèle est que la

volatilité σ de l'actif s dépend uniquement de la date et du prix de l'actif x. Nous verrons dans la suite que les autres paramètres de la diffusion n'ont que peu d'importance.
Pour couvrir l'option, l'agent peut utiliser les stratégies autofinancées de la forme :

$$Y_r^{t,s,x,y,\pi} = y + \int_t^r \pi_S^{t,s,x}(u) dS_u^{t,s,x} + \int_t^r \pi_X^{t,s,x}(u) \left(dX_u^{t,s,x} + \mu(u, X_u^{t,s,x})du\right)$$

Le problème posé est de trouver le plus petit prix de surcouverture d'une option payant $g(S_T)$ en T, avec g une fonction C^2 et bornée. Ce prix s'écrit de la manière suivante :

$$v\left(t,s,x\right) = \inf_{y\in\mathbb{R}} \left\{ y : Y_T^{t,s,x,y,\pi} \geq g\left(S_T^{t,s,x}\right) \text{ p.s. pour un } \pi \in \mathcal{A}_{t,s,x} \right\}$$

Nous ajoutons les hypothèses suivantes sur le payoff et les coefficients de la diffusion :

Hypothèse 1.1. *La fonction de payoff g est borné, C^2 et la fonction $x \to x^2 g''(x)$ est bornée.*
Les fonction σ^2 et μ sont localement Lipschitz et à croissance linéaire sur $(0,T) \times (0,+\infty)$.

Il reste à décrire l'ensemble $\mathcal{A}_{t,s,x}$ des stratégies admissibles. Nous cherchons à limiter les variations de la quantité d'actif x détenue par l'agent. D'une part car en pratique celui ci pourrait être soumis à des coûts de transaction, et d'autre part car cela permettra de s'affranchir de connaître précisément la dynamique de x. Pour cela, on impose, en utilisant les contraintes gamma de Cheridito, Soner et Touzi [23], que le processus décrivant la quantité d'actif x détenue par l'agent soit à variation finie. Les stratégies concernant l'actif x seront donc contraintes à être de la forme :

$$\pi_X(r) = \sum_{n=0}^{N-1} y_x^n 1_{\tau_x^n > t} + \int_t^r \alpha_x(u) du,$$

Les stratégies sur s sont elles non contraintes et peuvent être de la forme :

$$\pi_S(r) = \sum_{n=0}^{N-1} y_s^n 1_{\tau_s^n > t} + \int_t^r \alpha_s(u) du + \int_t^r \gamma^{s,s}(u) dS_u^{t,s} + \int_t^r \gamma^{s,x}(u) dX_u^{t,x},$$

où les τ_s^n, τ_x^n sont des temps d'arrêts croissants en n, les y_s^n, y_x^n sont respectivement $\mathcal{F}_{\tau_n^s}$ et $\mathcal{F}_{\tau_n^s}$ mesurables et les processus α_x, α_s sont adaptés et borné presque sûrement. Pour des raisons techniques, les processus $\gamma^{s,s}$ et $\gamma^{s,x}$ doivent pouvoir s'écrire de la même forme que π_S.
Les différences avec les hypothèses de Cheridito, Soner et Touzi [23] sont les suivantes :

- Dans leur article, la matrice γ est supposée être symétrique. Dans notre travail, nous ne le supposerons pas. On peut donc avoir, en théorie, $\gamma^{s,x} \neq \gamma^{x,s}$. Par ailleurs les contraintes ne sont pas les mêmes pour ces deux composantes. Nous donnons, en annexe, un résultat complétant celui de Cheridito, Soner et Touzi [24], permettant de voir que cette asymétrie n'influe pas sur le résultat.
- De plus, dans leur cadre, la matrice γ est contrainte d'évoluer dans un ensemble convexe d'intérieur non vide. Ici, la contrainte $\gamma^{s,x} = \gamma^{x,x} = 0$, rend l'intérieur de cet ensemble vide. Ceci modifie alors la forme de l'opérateur utilisé dans l'EDP, et introduit l'importance de la condition au bord en x.

On cherche dans cette partie à caractériser le prix de surréplication $v\,(t,s,x)$ par une équation aux dérivées partielles. La notion d'EDP utilisée ici est celle de solution de viscosité. On pourra se référer à l'article de Crandall, Ishii et Lions [27] pour une introduction à ce type de solutions. Pour décrire l'EDP dans notre cas on introduit l'opérateur :

$$F\left(t,s,x,Du,D^2u\right) = \lambda^- \begin{pmatrix} -\frac{\partial u}{\partial t} + \mu(t,x)\frac{\partial u}{\partial x} - \frac{1}{2}s^2\sigma^2(t,x)\frac{\partial^2 u}{\partial s^2} & -\frac{1}{2}s\sigma(t,x)\xi(t,x)\frac{\partial^2 u}{\partial s \partial x} \\ -\frac{1}{2}s\sigma(t,x)\xi(t,x)\frac{\partial^2 u}{\partial s\partial x} & -\frac{1}{2}\left(\xi(t,x)\right)^2\frac{\partial^2 u}{\partial x^2} \end{pmatrix},$$

où λ^- désigne la plus petite valeur propre de la matrice. Pour avoir une solution unique à l'EDP, il faut caractériser le prix aux bord du domaine en $t = T$ et en $x = 0$. On obtient les conditions aux bord :

$$\lim_{t \nearrow T, s' \to s, x' \to x} v(t,s',x') = g(s), \tag{1.1}$$

$$\lim_{t' \to t, s' \to s, x' \to 0} v(t',s',x') = g(s). \tag{1.2}$$

On peut ensuite caractériser le prix de surréplication, en montrant qu'il est solution de viscosité de l'EDP suivante, puis en prouvant l'unicité de la solution de l'équation grâce à un principe de comparaison.

Theoreme 1.1. *Le prix de surréplication v est l'unique solution de viscosité bornée de l'équation :*

$$F\left(t,s,x,Dv,D^2v\right) = 0 \tag{1.3}$$

sur le domaine $[0,T) \times [0,+\infty) \times [0,+\infty)$ vérifiant les conditions aux bord (1.1) *et* (1.2). *De plus v est continue sur son ensemble de définition.*

Nous pouvons par ailleurs obtenir une autre représentation de ce problème en utilisant l'EDP caractéristique. En effet, l'équation (1.3) peut se réecrire de manière équivalente :

$$-\frac{\partial v}{\partial x} - \sup_{\xi,\rho}\left\{-\frac{\partial v}{\partial x}\mu(t,x) + \frac{1}{2}\xi^2\frac{\partial^2 v}{\partial x^2} + \frac{1}{2}\sigma^2(t,x)s^2\frac{\partial^2 v}{\partial s^2} + \rho\xi\sigma(t,x)s\frac{\partial v}{\partial s\partial x}\right\} = 0.$$

On définit alors la diffusion contrôlée suivante :

$$\begin{aligned} S_t^{\rho,\xi,t,s,x} &= s \quad \text{et} \quad X_t^{\rho,\xi,t,s,x} = x \\ dS_u^{\rho,\xi,t,s,x} &= \sigma\left(t, X_u^{\rho,\xi,t,s,x}\right) S_u^{\rho,\xi,t,s,x} dW_u^1 \\ dX_u^{\rho,\xi,t,s,x} &= -\mu\left(t, X_u^{\rho,\xi,t,s,x}\right) du + \xi_u X_u^{\rho,\xi,t,s,x} dW_u^2 \\ \left\langle dW_u^1, dW_u^2\right\rangle &= \rho_u, \end{aligned}$$

où $(\xi,\rho) \in \mathcal{U}$ sont des contrôles. Leur ensemble admissible s'écrit :

$$\mathcal{U} = \left\{(\rho,\xi)\text{a valeur dans } [-1,1]\times[0,+\infty), \text{ adaptés, t.q : } \int_0^T \xi_t^2 dt < +\infty\right\}.$$

On obtient alors un résultat semblable à celui obtenu dans Cheridito, Soner et Touzi [24] :

Theoreme 1.2. *Le prix de surréplication v admet la représentation suivante :*

$$v(t,s,x) = \sup_{(\rho,\xi)\in\mathcal{U}} \mathbb{E}\left[g\left(S_{t,s,x}^{\rho,\xi}(T)\right)\right]. \tag{1.4}$$

Ce résultat montre que le choix des coefficients de la diffusion du processus X n'a aucun impact sur le prix. La surréplication dans ce cadre ne demande pas de connaître ceux ci. Ceci nous assure donc une robustesse de la stratégie de surcouverture, par rapport à des paramètres difficiles à mesurer. Néanmoins, pour être parfaitement rigoureux, il faudrait montrer que la stratégie de couverture ne dépend pas non plus de ces paramètres.

1.2 Chapitre 2 : Algorithme de résolution numérique

Ce chapitre a été réalisé en collaboration avec O. Bokanowski, S. Maroso et H. Zidani. Elle porte sur l'étude de la résolution numérique de l'équation obtenue dans le chapitre précédent. En effet, l'EDP (1.3) est sous une forme non standard. Pour résoudre ce type d'équation, il existe deux grandes familles de méthodes : les méthodes probabilistes, et les approches par EDP (différences finies, éléments finis). Nous utiliserons ici des méthodes par différences finies. Nous commençons par changer la forme de l'équation en remarquant que la plus petite valeur propre d'une matrice symétrique J s'écrit :

$$\Lambda^-(J) = \min_{\|\alpha\|=1} \alpha^T J\alpha,$$

pour $\alpha \in \mathbb{R}^2$. L'EDP (1.3) s'écrit alors :

$$\min_{\alpha_1^2+\alpha_2^2=1}\left\{ \begin{aligned} &-\alpha_1^2\frac{\partial v}{\partial t}(t,s,x) + \mu(t,x)\alpha_1^2\frac{\partial v}{\partial x}(t,s,x) \\ &-\frac{1}{2}\mathrm{tr}[a(\alpha_1,\alpha_2,t,s,x)D^2v(t,s,x)]\end{aligned}\right\} = 0, \tag{1.5}$$

avec :

$$a(\alpha_1,\alpha_2,t,s,x) \quad := \quad \begin{pmatrix} \alpha_1\sigma(t,x)s \\ \xi(t,x)\alpha_2 \end{pmatrix}\begin{pmatrix} \alpha_1\sigma(t,x)s \\ \alpha_2\xi(t,x) \end{pmatrix}^{\mathsf{T}}.$$

on remarquera en particulier le terme en α_1^2 devant le terme de dérivée en temps. Ce terme peut s'annuler, ce qui est la principale particularité de l'équation. Celle ci peut être interprétée comme la prise en compte des cas où $\xi \to \infty$ dans la formulation (1.4) de la fonction valeur comme problème de contrôle, tout en gardant des coefficients bornés dans l'EDP. Cette idée pourrait être adaptée à d'autres problèmes de contrôle non bornés.
L'autre difficulté du problème vient de la matrice a, de rang 1. En effet, d'après les travaux de Kushner [33], la consistance des schémas de différences finies classiques nécessite une matrice a à diagonale dominante, afin d'obtenir une interprétation probabiliste. Ici, cette condition n'est remplie que pour $\alpha_1 = 0$ ou $\alpha_2 = 0$. Nous utiliserons donc un autre type de schéma, fourni par Bonnans, Zidani et Ottenwaelter [18],[17]. Pour obtenir la consistance, la discrétisation en espace utilise non seulement les points immédiatement voisins du point considéré, mais aussi les points éloignés, afin prendre en compte un maximum de direction de diffusions dégénérées. On obtient grâce à cela un schéma consistant. Cette discrétisation prend la forme :

$$\Delta_\zeta\phi(t,x,y) = \phi(t,x+\zeta_1h_1,y+\zeta_2h_2) + \phi(t,x-\zeta_1h_1,y-\zeta_2h_2) - 2\phi(t,x,y)$$

où $\zeta = (\zeta_1,\zeta_2) \in \mathbb{N}^2$ est la direction de la diffusion, et (h_1,h_2) sont les pas de discrétisation pour chaque coordonnée spatiale. En ce qui concerne la discrétisation en temps, on utilisera un schéma implicite. En posant $\rho = (p, h, \Delta t) \in \mathbb{N} \times (0, +\infty)^2$ on obtient ainsi un schéma de la forme générale :

$$S^\rho(t,s,x,r,\phi) \quad = \quad \min_{\alpha_1^2+\alpha_2^2=1} \Big\{ -\alpha_1^2 \frac{\phi(t+\Delta t,s,x)-r}{\Delta t} + \alpha_1^2 \mu \frac{r-\phi(t,s,x-h)}{h}$$
$$-\tfrac{1}{2}\sum_{\zeta\in\{0,..,p\}^2} \gamma_\zeta^{\alpha_1,\alpha_2}(t,s,x)\big[\phi(t,s-\zeta_1 h,x-\zeta_2 h)-2r+\phi(t,s+\zeta_1 h,x+\zeta_2 h)\big]\Big\}.$$

Les coefficients γ ne seront pas explicités dans cette introduction. Tout au moins, nous pouvons dire qu'il y en a uniquement trois non nuls.
On prouvera que ce schéma satisfait les trois hypothèses principales assurant la convergence :

Proposition 1.1. *Le schéma* (1.6) *satisfait les hypothèses classiques :*

(S1) ***Monotonie :*** $S^\rho(t,x,y,r,u) \geq S^\rho(t,x,y,r,v)$,
pour tout $r \in \mathbb{R}$, $x, y \in \mathbb{R}_+^*$, $u, v \in C([0,T]\times[0,\infty)^2)$ *tel que* $u \leq v$ *in* $[0,T]\times[0,\infty)^2$.

(S2) ***Stabilité :*** *Pour tout* $\rho = (h, \Delta t, p_{\max}) \in (\mathbb{R}_+^*) \times (0,T) \times \mathbb{N}^*$, *il existe une solution bornée* v_h *de (1.6).*

(S3) ***Consistence :*** *Pour toute fonction* $\phi \in C^n([0,T]\times[0,\infty)^2)$, $n \geq 4$, *à dérivées bornées, si* $p = o(\frac{1}{h})$, *et il existe* C *t.q.* $p_{max} \geq \frac{C}{\sqrt{h}}$ *alors*

$$\Bigg| \min_{\alpha_1^2+\alpha_2^2} \Big\{ -\alpha_1^2 \frac{\partial \phi}{\partial t}(t,s,x) + \alpha_1^2 \mu \frac{\partial \phi}{\partial y}(t,s,x) - \frac{1}{2}\mathrm{tr}[a \cdot D^2\phi(t,s,x)]\Big\}$$
$$-S^\rho(t,s,x,\phi(t,s,x),\phi)\Bigg| = O(h) + O(\Delta t),$$

pour tout $(t,s,x) \in [0,T] \times (\mathbb{R}_+^*)^2$.

Il faut encore montrer l'existence d'une solution au schéma. Comme on considère un domaine infini, nous montrons tout d'abord, à pas de discrétisation ρ et contrôle α fixé, l'existence d'une unique solution bornée au système linéaire infini décrit par le schéma implicite. Ensuite, nous utilisons l'algorithme de Howard [46] pour obtenir une suite convergeant simplement vers la solution du schéma discret (1.6).
En utilisant la proposition 1.1, ainsi que les résultats de Barles et Souganidis [10] on est alors en mesure de prouver la convergence de l'algorithme. On obtient finalement le théorème suivant :

Theoreme 1.3. *Sous les hypothèses 1.1, si p satisfait les hypothèses de la proposition* (1.1) *(S3), alors la solution du schéma discret converge localement uniformément vers* v *quand* $h, \Delta t \to 0$.

2 Présentation des résultats de la deuxième partie

2.1 Chapitre 3 : Valorisation d'option avec volatilité incertaine et tolérance aux pertes

Cette partie concerne la valorisation et couverture d'options Européennes en présence de risque de volatilité. Nous nous placons dans le cadre d'un marché comprenant $d+1$ actifs financiers. On considère les prix actualisés. Celui de l'actif sans risque sera donc supposé constant et égal à 1. Nous ne supposerons aucune dynamique particulière pour la matrice de volatilité. Les prix des actif risqués évoluent de la manière suivante :

$$dS_t^\sigma = diag(S^{\sigma_t})\sigma_t dW_t$$

Où W_t est un mouvement Brownien standard de dimension d et $\sigma \in \Sigma$ est un processus de dimension $d \times d$ dont les caractéristiques seront données dans la suite. On cherche à valoriser une option dont le payoff est $g(S_T)$ en T, avec g une fonction continue et bornée. Si on valorise une option dans le modèle de Black et Scholes avec une volatilité $\hat{\sigma}$, on obtient un prix $P^{BS}(t,s)$. On peut alors estimer les pertes potentielles par les formules de robustesse de El Karoui, Jeanblanc et Shreve [35]. On obtient alors, dans un marché à un seul actif risqué, une erreur de couverture Y_T de la forme :

$$Y_T = \int_0^T \frac{1}{2} S_t^2 \frac{\partial^2 P^{BS}}{\partial t} \left(\hat{\sigma}^2 - \sigma_t^2\right) dt.$$

Cette formule a connu un grand succès parmi les praticiens, du fait de sa simplicité et de la relation explicite entre la volatilité réalisée et les profits. L'idée dévelopée dans notre travail est d'effectuer la démarche inverse. Nous allons partir d'une formule de robustesse pour aboutir au prix de l'option.
Lorsque l'on ne tolère aucune perte, (c.a.d $Y_T \geq 0$ p.s.), des réponses ont été apportées par Cvitanic,Pham et Touzi [29] puis par Avellaneda,Levy,Paras [3] et Gozzi,Vargiolu [40]. Les résultats dépendent alors de l'ensemble dans lequel la volatilité est susceptible d'évoluer. Lorsque celle ci n'est pas bornée, on obtient des prix de surréplication correspondant à des stratégies de couverture statique et triviales. En revanche, lorsque l'on suppose la volatilité bornée, on obtient des prix non triviaux. Le problème est que si jamais la volatilité ne vérifie pas ces bornes, on a alors aucun contrôle sur l'erreur de couverture de l'option. Nous cherchons donc à combiner les avantages de ces deux points de vue.
Nous considèrerons donc une volatilité non bornée a priori, et nous admettrons des pertes de la forme :

$$\int_0^T f(t, S_t, \sigma_t^2) dt.$$

où $f : [0,T] \times \mathbb{R}^d \times \mathcal{S}_d \to \mathbb{R} \cup \{+\infty\}$. Le prix de l'option avec cette tolérance aux pertes s'écrira donc :

$$\begin{aligned} v(t,s) = \inf_z \Big\{ & z \in \mathbb{R} : \text{ il existe } \pi \in \mathcal{A}_1 \text{ tel que} \\ & z + \int_t^T \pi_u dS_u^{t,s,\sigma} \geq g(S_T^{t,s,\sigma}) - \int_t^T f(u, S_u^{t,s,\sigma}, \sigma_u^2) du \text{ p.s. pour tout } \sigma \in \Sigma \Big\} \end{aligned}$$

Ici, $\mathcal{A}_1$ est l'ensemble des processus adaptés tels que $\int_0^T \pi_t dS_u^{0,s,\sigma}$ est bornée presque surement pour tout $\sigma \in \Sigma$. Σ est défini par :

$$\begin{aligned} \Sigma = \Big\{ & \sigma = (\sigma_t)_{t \in [0,T]} \text{ processus adapté, borné p.s., à valeurs dans } \mathcal{S}_+^d \\ & \text{t.q. } \int_0^T f(t, S_t^{0,s,\sigma}, \sigma_t) dt \text{ est borné p.s. } \Big\}. \end{aligned}$$

Si on tolère des pertes infinie lorsque la volatilité reste dans un certain ensemble et des pertes nulles sinon, on retrouve alors les résultats de [40] et [3]. Pour caractériser le prix ainsi défini, nous introduisons tout d'abord la transformée de Fenchel $\tilde{f}$ de la fonction f :

$$\tilde{f}(t,s,A) = \sup_{\sigma^2} \left\{ \frac{1}{2} Tr\left(A\sigma^2\right) - f(t,s,\sigma^2) \right\}.$$

Nous introduisons ensuite les hypothèses techniques suivantes :

Hypothèse 2.1. *(i) Pour tout $(t,s) \in [0,T] \times \mathbb{R}^d$, la fonction :*

$$\sigma^2 \to f(t,s,\sigma^2)$$

est convexe et semi continue inférieurement.
(ii) La fonction f est continue par rapport à (t,s) uniformément en (t,s,σ^2).
(iii) La fonction f est bornée inférieurement, et il existe une fonction bornée :

$$\begin{aligned} \sigma : (0,+\infty)^d &\to \mathcal{S}_+^d \\ (t,s) &\to \sigma(t,s) \end{aligned}$$

et une constante C telle que $f(t,s,\sigma^2(t,s)) < C$, pour tout $(t,s) \in [O,T] \times (0,+\infty)$.
(iv) Pour tout $\varepsilon > 0$, il existe K_ε tel que :

$$\tilde{f} \ est \ K_\varepsilon - Lip \ dans \ int_\varepsilon(dom\left(\tilde{f}(t,s,.)\right))\forall(t,s)$$

et $0 \in int\left(dom(\tilde{f}(t,s,.))\right) \forall(t,s)$

D'après la forme donnée au prix, le contexte naturel serait d'utiliser le cadre de surréplication de Soner et Touzi [64]. Cependant, le principe de la programmation dynamique pour la sur-couverture avec incertitude sur la volatilité a été uniquement prouvé dans le cas borné par Denis et Martini [32]. Pour éviter ce problème, nous introduisons donc le problème de contrôle standard :

$$w(t,s) = \sup_{\sigma \in \Sigma} \mathbb{E}[g(S_T^{t,s,\sigma}) - \int_t^T f(u, S_u^{t,s,\sigma}, \sigma_u^2) du],$$

qui, d'après les travaux de El Karoui et Quenez [36] dans un contexte légèrement différent, devrait avoir la même valeur que le prix v. Nous utiliserons la caractérisation du prix par EDP pour montrer cela. Nous introduisons d'abord l'opérateur G défini par :

$$G(t,s,A) = \begin{cases} \inf\left\{|B| \text{ t.q. } A+B \notin dom(\tilde{f}(t,s,.))\right\} \text{ si } A \in dom(\tilde{f}(t,s,.)) \\ -\inf\left\{|B| \text{ t.q. } A+B \in dom(\tilde{f}(t,s,.))\right\} \text{ si } A \notin dom(\tilde{f}(t,s,.)) \end{cases}.$$

Grâce aux hypothèses formulées, on peut prouver que G ne dépend pas de (t,s). On le note donc $G(A)$. On définit enfin l'opérateur :

$$\begin{aligned} F(t,s,p,B) = \sup_{A \geq 0} \Big\{ \min \Big\{ & -p - \tilde{f}\left(t,s,\mathrm{diag}[s]B\mathrm{diag}[s] - A\right), \\ & \mathbf{1}_{A=0} G\left(\mathrm{diag}[s]B\mathrm{diag}[s]\right) - tr\left(A\right)\Big\}\Big\}. \end{aligned} \tag{2.1}$$

L'opérateur a cette forme particulière afin de pouvoir contrôler son comportement lorsque σ tend vers l'infini, ce qui pourrait entrainer des valeurs infinies du Hamiltonien écrit sous forme standard. On peut alors prouver le théorème suivant, caractérisant la fonction w.

Theoreme 2.1. *Si les hypothèses précédentes sont vérifiées, alors w est continue et est l'unique solution de viscosité bornée de l'équation :*

$$F(t, s, \frac{\partial w}{\partial t}, D_s^2 w) = 0$$

satisfaisant la condition terminale $w(T^-, .) = \hat{g}$, où $\hat{g}$ est caractérisée comme l'unique solution de viscosité de l'équation :

$$\min\left\{\hat{g}(s) - g(s), G(diag[s]\, D^2\hat{g} diag[s])\right\} = 0. \tag{2.2}$$

Il reste enfin à prouver que la fonction w est bien celle que l'on cherche. La preuve que $v \geq w$ est quasiment immédiate, en utilisant un argument de surmartingale. En revanche, la preuve de $v \leq w$ est basée sur un argument de couverture faisant appel à l'EDP vérifiée par w, et à des procédures de régularisation de cette fonction pour pouvoir lui appliquer la formule d'Itô. On obtient finalement :

Theoreme 2.2. *Les fonctions valeur des deux problèmes sont égales :*

$$v = w$$

3 Présention des résultats de la troisième partie

3.1 Chapitre 4 : Problème de controle optimal avec retard à l'execution

Dans cette troisième partie, réalisée en collaboration avec Huyen Pham, nous considérons un problème de contrôle pour lequel les actions de l'agent prennent effet avec retard. Cette étude a été motivée par le problème de couverture d'options dont le sous jacent est un hedge fund. Ce type de fonds ayant récemment connu un succès grandissant auprès des investisseurs, les banques commencent à les utiliser comme sous jaçent des options qu'elles émettent. Cependant, ceux ci posent des problèmes de liquidité. En effet, les parts de hedge funds sont des actifs qui ne s'échangent pas au sein de marchés organisés. Au contraire, elles sont achetées et vendues directement auprès du gérant du fonds. Celui ci crée donc de nouvelles parts lors des demandes d'achat, et les liquide lorsque des ordres de ventes sont passés. Or, la plupart de ces fonds investissent dans des actifs eux mêmes illiquides. Du temps est alors nécessaire pour trouver de nouvelles opportunité d'investissement, ou pour liquider une partie de leurs actifs. La solution pour le gérant du hedge fund est alors d'imposer aux autres agents de déclarer leurs ordres d'achat ou de ventes de parts un ou plusieurs mois à l'avance. Bien sûr, une fois l'ordre déclaré, il n'est plus possible de l'annuler. Le prix auquel l'agent effectuera la transaction sera celui de la part au moment de l'exécution et non du passage d'ordres.
Lors de la couverture d'une option sur un tel sous jaçent, les ordres passés par l'agent subissent donc un retard d'un ou plusieurs mois. La couverture sera donc imparfaite, et il convient d'utiliser un critère tel le prix par indifférence de Hodges [45].
Pour pouvoir le mettre en oeuvre, nous avons choisi d'étudier de manière générale un classe de problème de contrôles impulsionnel avec retard. Pour une introduction au contrôle impulsionnel, on pourra se réferrer au livre de Bensoussan et Lions [13]. Dans notre problème, nous considérons un processus X, d-dimensionnel. Lorsque l'agent n'agit pas, sa dynamique est données par l'EDS suivante :

$$dX_s \quad = \quad b(X_s)ds + \sigma(X_s)dW_s,$$

où W_s est un mouvement Brownien standard de dimension n, et $b : \mathbb{R}^d \to \mathbb{R}^d$ et $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times n}$ satisfont les conditions de Lipschitz usuelles.
A tout temps d'arrêt τ_i, l'agent peut décider d'agir sur le système en passant une impulsion ξ_i basée sur l'information disponible en τ_i. Cependant, cette impulsion prendra effet avec retard en $\tau_i + mh$, où $(m, h) \in \mathbb{N} \times (0, +\infty)$ sont des constantes. Le système subira alors l'évolution suivante :

$$X_{(\tau_i+mh)} = \Gamma(X_{(\tau_i+mh)^-}, \xi_i).$$

On suppose que la fonction Γ est continue, et satisfait une condition de croissance linéaire. L'ensemble des contrôles admissibles s'écrit :

$$\mathcal{A} \quad = \quad \Big\{\alpha = (\tau_i, \xi_i)_{i\geq 1} : \tau_i \text{ est un t.a., } \xi_i \text{ est } \mathcal{F}_{\tau_i}\text{adapté}, \tau_{i+1} - \tau_i \geq h\Big\}.$$

On constate ci dessus que l'on impose un temps minimal h entre deux interventions de l'agent. En se donnant un contrôle $\alpha \in \mathcal{A}$, et une condition initiale $X_0 \in \mathbb{R}^d$, le processus contrôlé X^α est alors solution de l'EDS :

$$X_s = X_0 + \int_0^s b(X_u)du + \int_0^s \sigma(X_u)dW_u + \sum_{\tau_i+mh\leq s} \big(\Gamma(X_{(\tau_i+mh)^-}, \xi_i) - X_{(\tau_i+mh)^-}\big).$$

Le but du problème est de maximiser l'espérance du gain de l'agent en T. Ceci s'écrit :

$$V_0 = \sup_{\alpha\in\mathcal{A}} \mathbb{E}\left[\int_0^T f(X_s^\alpha)ds + g(X_T^\alpha) + \sum_{\tau_i+mh\leq T} c(X^\alpha_{(\tau_i+mh)^-}, \xi_i)\right].$$

On suppose que les fonctions de gain terminal g, de gain continu f et de gains aux transactions c sont continues et satisfont une condition de croissance linéaire. Nous supposons aussi, afin de pouvoir obtenir la continuité de la fonction valeur :

Hypothèse 3.1. *Les fonctions Γ, g et c vérifient :*

$$g(x) \geq g\big(\Gamma(x, e)\big) + c(x, e)$$

pour tout $(x, e) \in \mathbb{R}^d \times E$.

Cela entraine qu'il n'est jamais optimal de passer une impulsion en $T - mh$.
Nous cherchons ensuite a obtenir un système Markovien. Pour cela, nous devons introduire une variable supplémentaire, p, représentant les ordres passés mais non encore exécutés. Notons que grâce à notre contrainte imposant un temps minimal h entre deux interventions, il y a au plus m ordres en attente. A une date $t \in [0, T]$, si il y a $k \in \{0..m\}$ ordres en attente, cette variable p appartient à l'ensemble :

$$\begin{aligned} P_t(k) \quad = \quad \Big\{ & p = (t_i, e_i)_{1\leq i\leq k} \in ([0, T - mh] \times E)^k \ : t_i - t_{i-1} \geq h, \ \ i = 2, \dots, k, \\ & \qquad\qquad t - mh < t_i \leq t, \ i = 1, \dots, k\Big\}. \end{aligned}$$

On peut ensuite définir les contrôles admissibles à partir d'une date t et d'un ensemble d'ordre en attentes p :

$$\mathcal{A}_{t,p} \quad = \quad \Big\{\alpha = (\tau_i, \xi_i)_{i\geq 1} \in \mathcal{A} \ : (\tau_i, \xi_i) = (t_i, e_i), \ i = 1, \dots, k \ \text{ and } \ \tau_{k+1} \geq t\Big\}.$$

Ceci permet alors de définir la fonction valeur du problème démarrant en (t, x, p) :

$$v_k(t, x, p) = \sup_{\alpha \in \mathcal{A}} \mathbb{E}\left[\int_t^T f(X_s^{\alpha,t,x})ds + g(X_T^{\alpha,t,x}) + \sum_{\tau_i + mh \le T} c(X_{(\tau_i+mh)^-}^{\alpha,t,x}, \xi_i)\right],$$

le processus X admettant x comme condition initiale en t. Enfin, on note $\mathcal{D}_k$ le domaine de définition de v_k :

$$\mathcal{D}_k \;=\; \big\{(t, x, p) \;:(t, x) \in [0, T] \times \mathbb{R}^d,\ p \in P_t(k)\big\},$$

que l'on partitionne en deux sous ensembles :

$$\begin{aligned}
\mathcal{D}_k^1 &= \big\{(t, x, p) \;:(t, x) \in [0, T] \times \mathbb{R}^d,\ p \in P_t(k) \text{ t.q. } t_k + h > t \text{ ou } t > T - mh\big\} \\
\mathcal{D}_k^2 &= \big\{(t, x, p) \;:(t, x) \in [0, T] \times \mathbb{R}^d,\ p \in P_t(k) \text{ t.q. } t_k + h \le t \text{ et } t \ge T - mh\big\}
\end{aligned}$$

Le théorème suivant énonce le principe de la programmation dynamique utilisé pour caractériser la fonction valeur.

Theoreme 3.1. *On note :*

$$\begin{aligned}
\iota(t, \alpha) &= \inf\{i \ge 1 \;: \tau_i > t - mh\} - 1 \;\in\; \mathbb{N} \cup \{\infty\}, \\
k(t, \alpha) &= \operatorname{card}\big\{i \ge 1 \;: t - mh < \tau_i \le t\big\} \;\in\; \{0, \ldots, m\}, \\
p(t, \alpha) &= (\tau_{i+\iota(t,\alpha)}, \xi_{i+\iota(t,\alpha)})_{1 \le i \le k(t,\alpha)} \;\in\; P_t(k(t, \alpha)).
\end{aligned}$$

La fonction valeur satisfait alors le principe de programmation dynamique suivant : pour tout $k = 0, \ldots, m$, $(t, x, p) \in \mathcal{D}_k$,

$$\begin{aligned}
v_k(t, x, p) \;=\; \sup_{\alpha \in \mathcal{A}_{t,p}} \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha})ds &+ \sum_{\tau_i + mh \le \theta} c(X_{(\tau_i+mh)^-}^{t,x,p,\alpha}, \xi_i) \\
&+ v_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta, \alpha))\Big],
\end{aligned} \tag{3.1}$$

pour tout temps d'arrêt θ à valeurs dans $[t, T]$, dépendant éventuellement de α dans (3.1).

Ce principe de la programmation dynamique nous permet alors de montrer la caractérisation des fonctions valeurs en termes d'EDP. Tout d'abord nous donnons les conditions terminales :

Proposition 3.1. *(i) Pour $k = 1, \ldots, m$, $p = (t_i, e_i)_{1 \le i \le k} \in \Theta_k^m \times E^k$, $x \in \mathbb{R}^d$, $v_k((t_1 + mh)^-, x, p)$ existe et :*

$$v_k((t_1 + mh)^-, x, p) \;=\; c(x, e_1) + v_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-). \tag{3.2}$$

(ii) Pour $k = 1, \ldots, m$, on a :

$$v_k(t, x, p) \;=\; \mathbb{E}\Big[\int_t^T f(X_s^{t,x,0})ds + g(X_T^{t,x,0})\Big], \tag{3.3}$$

pour tout $(t, x, p) \in \mathcal{D}_k$ t.q. $t_1 + mh > T$.

En utilisant le principe de programmation dynamique pour montrer que la fonction vérifie l'EDP, puis un théorème de comparaison pour montrer l'unicité de la solution sachant les conditions au bord, on obtient le résultat principal.

Theoreme 3.2. *La famille de fonctions valeurs v_k, $k = 0, \ldots, m$, est l'unique solution de viscosité des équations*

$$-\frac{\partial v_k}{\partial t}(t,x,p) - \mathcal{L}v_k(t,x,p) - f(x) \;=\; 0 \quad sur \quad \mathcal{D}_k^{1,m}, \quad k = 1, \ldots, m, \tag{3.4}$$

$$\min\Big\{ \; -\frac{\partial v_k}{\partial t}(t,x,p) - \mathcal{L}v_k(t,x,p) - f(x)\;, \quad v_k(t,x,p) - \sup_{e\in E} v_{k+1}(t,x,p\cup(t,e))\Big\} \;=\; 0 \;\; sur \;\; \mathcal{D}_k^{2,m}, \;\; k = 0, \ldots, m-1 \tag{3.5}$$

satisfaisant les conditions aux bords (3.2)-(3.3)*, une condition de croissance linéaire, et l'inégalité :*

$$\underline{v_k}(t,x,p) \geq \sup_{e\in E} \overline{v_{k+1}}(t,x,p\cup(t,e)),$$

pour tout $k = 0, \ldots, m-1$, $(t,x,p) \in \mathcal{D}_k^2$, $p = (t_i,e_i)_{1\leq i\leq k}$, tel que $t = t_k + h$ ou $t = T - mh$. De plus, v_k est continue sur $\mathcal{D}_k$, $k = 0, \ldots, m$.

3.2 Chapitre 5 : Résolution numérique du problème de controle avec retard

Dans ce chapitre, nous cherchons à approximer numériquement les fonctions valeurs v_k, $k = 0, \ldots, m$ décrites dans le problème de controle impulsionnel avec retard du chapitre 4. Les évaluations numériques de problèmes de contrôle impulsionnel ont fait l'objet de nombreux travaux. On pourra, pour une introduction générale, se référer au livre d'Oksendal et Sulem [56], ainsi qu'aux travaux récents de Chen et Forsyth [22]. D'autres problèmes proches, issus du contrôle singulier sont traités par Hodder, Tourin et Zariphopoulou [44]. Ici, nous résoudrons numériquement l'EDP linéaire (3.4) et l'inéquation variationelle (3.5). La première équation étant linéaire, sa discrétisation rentre dans le cadre standard. La deuxième équation peut être interprétée comme un problème d'arrêt optimal, étudié notamment dans les travaux de Barles et Daher [8]. On peut donc utiliser les discrétisations standard pour ce type de problème en considérant $\sup_{e\in E} v_{k+1}$ comme un obstacle donné. La difficulté, ici, est due au fait que la condition terminale (3.2) de ces deux équations, ainsi que l'obstacle v_{k+1} sont endogènes au problème et doivent être eux mêmes calculés numériquement. Le premier apport de ce travail est de fournir un algorithme décrivant l'ordre dans lequel il est possible de calculer les fonctions v_k, $k = 1 \ldots m$. Nous introduisons pour cela les ensembles suivants :

$$\begin{aligned}
\mathcal{D}_k(n) &= \{(t,x,p) \in \mathcal{D}_k \;:\; t_1 > T - nh\} \\
\mathcal{D}_k^1(n) &= \mathcal{D}_k^1 \cap \mathcal{D}_k(n) \\
\mathcal{D}_k^2(n) &= \mathcal{D}_k^2 \cap \mathcal{D}_k(n)
\end{aligned}$$

– Tout d'abord, on remarque que la fonction valeur v_0 satisfait l'équation linéaire (3.4) sur l'ensemble $\mathcal{D}_0(m)$. Grâce à la condition terminale (3.3), on peut alors calculer v_0 sur cet ensemble.

– Ensuite, on procède par récurrence croissante sur $n = m, \dots, N$ et, à chaque pas sur n, par récurrence décroissante sur $k = m(n), \dots, 0$, avec $m(n) = (n-m) \wedge m$ étant la plus grande valeur de k telle que $\mathcal{D}_k(n)$ est non vide. Supposons que l'on connaisse les fonctions valeurs v_k sur les ensembles $\mathcal{D}_k(n-1)$ pour $k = 0, \dots, m(n-1)$. On remarque alors que $v_{m(n)}$ satisfait l'équation linéaire (3.4) sur $\mathcal{D}_{m(n)}(n)/\mathcal{D}_{m(n)}(n-1)$. De plus la condition terminale de l'équation fait appel aux valeurs de $v_{m(n)-1}$ sur $\mathcal{D}_{m(n)-1}(n-1)$, qui ont déjà été calculées à une étape précédente. On peut donc calculer la fonction valeur sur $\mathcal{D}_{m(n)}(n)$.
– Enfin, on suppose que l'on connaît, pour un certain $k \in \{0, \dots, m(n)-1\}$, la valeur de v_{k+1} sur $\mathcal{D}_{k+1}(n)$. On connaît alors l'obstacle dans l'équation (3.5). On peut alors calculer la fonction v_k sur sur $\mathcal{D}_k(n)$ grâce aux équations (3.4) et (3.5). En effet, la condition terminale de ces équations a déjà été calculée car elle fait appel aux valeurs de v_{k-1} sur $\mathcal{D}_{(k-1)}(n-1)$.

Ainsi, on peut calculer par récurrence les fonctions valeurs sur les domaines $\mathcal{D}_k$, $k = 0, \dots, m$, en supposant que l'on sait résoudre l'EDP linéaire et les inéquations variationnelles avec obstacle donné. Dans la suite du chapitre, nous donnons un schéma numérique pour résoudre ces types d'équations. On se donne un vecteur δ représentant le pas de discrétisation dans chaque direction de l'espace. Sur l'ensemble $\mathcal{D}_k^1$, $k = 0, \dots, m$, nous avons le schéma implicite correspondant à l'équation linéaire (3.4) :

$$\begin{aligned} \frac{S^{1,\delta}((t,x,p),r,\Psi_k)}{\delta_t} &= \\ \frac{r - \Psi_k(t+\delta_t,x,p)}{\delta_t} & \quad - L^\delta(t,x,p,r,\Psi_k), \end{aligned}$$

Où L^δ représente la discrétisation standard du générateur infinitésimal de la diffusion pour un schéma implicite. On pourra pour cela se référer par exemple au livre de Lapeyre, Sulem et Talay [49]. Sur l'ensemble $\mathcal{D}_k^2$, $k = 0, \dots, m-1$, le schéma correspondant à l'inéquation variationnelle (3.5) peut quand à lui s'écrire :

$$\begin{aligned} \frac{S^{2,\delta}((t,x,p),r,\Psi_k,\Psi_{k+1})}{\delta_t} &= \\ \min\Big\{ \frac{r - \Psi_k(t+\delta_t,x,p)}{\delta_t} &- L^\delta(t,x,p,r,\Psi_k), \\ \Psi_k(t,x,p) &- \sup_{e \in E^{\delta_e}} \{\Psi_{k+1}(t,x,p \cup (t,e))\} \Big\}. \end{aligned}$$

La non linéarité due au minimum ci dessus peut se résoudre grâce à l'algorithme de Howard. On peut alors définir le schéma numérique sur $\mathcal{D}_k$, $k = 0, \dots, m$:

$$\begin{aligned} S^\delta((t,x,p),r,\Psi_k,\Psi_{k+1}) &= 1_{(t,x,p)\in\mathcal{D}_k^1} S^{1,\delta}((t,x,p),r,\Psi_k) \\ &\quad + 1_{(t,x,p)\in\mathcal{D}_k^2} S^{2,\delta}((t,x,p),r,\Psi_k,\Psi_{k+1}) \end{aligned}$$

L'équation discrète à résoudre est alors, pour un pas de discrétisation δ :

$$S^\delta((t,x,p),\Psi_k(t,x,p),\Psi_k,\Psi_{k+1}) = 0, \tag{3.6}$$

sur $\mathcal{D}_k$, pour $k = 0, \dots, m$. Nous prouvons alors la stabilité, la monotonie et la consistance de ce schéma. Ceci nous permet, en utilisant la méthode de Barles et Souganidis [10] avec le principe de comparaison du chapitre précédent, de montrer la convergence du schéma.

Theoreme 3.3. *Pour tout pas de discrétisation δ, soit Ψ_k^δ la famille de solutions de* (3.6) *sur $\mathcal{D}_k$, $k = 0, \ldots, m$ satisfaisant les conditions terminales* (3.2) *et* (3.3)*. Alors pour tout $k = 0, \ldots, m$, Ψ_k^δ converge localement uniformément vers v_k sur $\mathcal{D}_k$ lorsque $\delta \to 0$.*

Enfin, dans une dernière partie, nous donnons un exemple d'application financière. Nous considérons un marché composé d'un actif sans risque considéré comme le numéraire, et d'un actif risqué dont le prix suit un processus S_t. Nous supposons que ce processus suit une modèle de Black-Scholes :

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t.$$

Nous modélisons le processus d'état par $X_t = (S_t, Y_t, Z_t)$, où Y_t représente le nombre de parts d'actif risqué détenu par l'agent et Z_t représente la quantité d'actif sans risque dans son portefeuille. Une impulsion ξ_i passée par l'agent à une date τ_i représentera la quantité d'actif risqué qu'il souhaite détenir à l'instant $\tau_i + mh$. Si $\xi_i > Y_{(\tau_i+mh)^-}$, cela représentera un ordre d'achat, et si $\xi_i < Y_{(\tau_i+mh)^-}$, cela représentera un ordre de vente. La fonction Γ représentant l'état du système après l'exécution de l'ordre pourra donc s'écrire :

$$\Gamma\Big(\begin{pmatrix} s \\ y \\ z \end{pmatrix}, e\Big) = \begin{pmatrix} s \\ e \\ z + (y-e)s \end{pmatrix}$$

Nous supposerons que l'agent possède une fonction d'utilité U portant sur la valeur liquidative $Z_T + S_T Y_T$ de son portefeuille à la date T. Dans ce cadre, nous donnerons des résultats numériques pour le problème de maximisation de l'espérance d'utilité de l'agent, puis pour celui du pricing par indifférence d'une option Européenne.

Part I

A super replication problem with gamma constraints

Chapter 1

Super-replication of European options with a derivative asset under constrained finite variation strategies

We consider a financial market, in which a first asset will be referred as the underlying and the second one as a derivative. In this market, the volatility on the underlying depends of the price of the derivative. Furthermore, the derivative is constrained to be traded with finite variation strategies. We study the super-replication problem of an European option on the underlying, and characterize its price as the unique viscosity solution of a partial differential equation with appropriate boundary conditions. We also give a dual representation of the price, as the supremum of the risk neutral expectation over a range of dynamics of the price of the derivative.

Key words : Gamma constraints, super-replication, viscosity solutions, double stochastic integrals

1 Introduction

It is commonly known that, under unbounded stochastic volatility, with no instrument to hedge oneself against this volatility, the super-replication price of an European option is the price of the cheapest buy and hold strategy involving the underlying. Hence this price is the concave envelope of the payoff of this option. This was treated, for example in [29]. Meanwhile, another problem gives the same result: the super replication price under constant volatility with fixed or proportional transaction costs. For example, see [12] or [28]. On the other hand, we know that in some stochastic volatility models, for example Heston's model [43] or Hull and White's model [47], one can perform a perfect hedge with the underlying and another instrument, used to hedge the volatility of the underlying. This is the case, for example, if another European option is traded on the same underlying. Nowadays, a new instrument tends to become the reference asset to hedge volatility: the variance swap. In this kind of contracts, a fixed payment is exchanged against the realized volatility of the asset. The most famous variance swap price index is the VIX index, which refers to the S&P500 American index. This kind of product has the benefit of simplifying lots of calculations compared to the call options. It can also merge all positions of the investors with respect to volatility in a single instrument, rather than on a market with call options of numerous strikes. Nevertheless, either the call options and variance swaps can be very illiquid and introduce lots of transaction costs. This is why, here, we will constraint these volatility hedging instruments to be traded with finite variations. These type of constraints are studied over the underlying in [12], and the result is again the cheapest buy and hold super-replicating strategy. But the case of constraints over the "volatility asset" with no constraints over the underlying is not yet considered in the literature, although important in practice. In this paper we focus on that case, and prove that the super-replication strategy is not necessarily a "Buy and Hold" strategy. Indeed, the superreplication price has to be concave with respect to the volatility asset price, but not w.r.t. the underlying. We characterize this price as the unique solution to a PDE in the viscosity sense, and the terminal condition is found to be the payoff itself. Moreover, we prove a dual representation as in [12] and [65], in which the price is the supremum of the risk neutral prices over all possible dynamics of the "volatility asset". Here, we do not consider vanishing transaction costs, but we require the quantity of asset in the portfolio to be almost surely of bounded variation, as a limit case of gamma constraints used in [23], when the authorized "gamma" with respect to one asset is zero. This is a new feature of this paper. Moreover, the gamma constraints considered here are not symmetric, which involve a new result about double stochastic integral as in [24], which is valid for non-symmetric integrands.

The structure of this paper is the following: In section 2, we define the model, the super-replication problem and the portfolio gamma constraints. We also state the main results. Then, in section 3 we show that the super-replication price is a solution of a partial differential equation with specific terminal and boundary conditions. The uniqueness of this solution is proved in section 4 with the help of a comparison principle. Finally, in section 5, we prove a dual representation of the solution, with can be interpreted as the supremum of the risk neutral prices of the option over a range of dynamics of the volatility asset.

2 Problem formulation and main results

2.1 Model

We consider a financial market with three different assets. The first one is a riskless bond, which we take as numeraire, so that the interest rate can be considered constant and equal to zero. The second one is a risky asset S, and the third one is an asset X whose price is linked to the instantaneous volatility $\sigma(t, X)$ of S. Note that it implies that the instantaneous volatility of the underlying is a given function of the price of a single instrument, which is not confirmed by statistical studies, see [26] for instance. This asset X distributes an instantaneous cash flow $\mu(t, X) \geq 0$. Indeed, these cash flows are typically positive, if we consider, for example X as a variance swap, for which the fixed leg would be payed upfront. Then we would have $\mu(t, X)=\sigma^2(t, X)$. Our problem is to find a super-replication price, hence we are only interested in almost sure events. Therefore we can specify our market under a risk neutral probability measure. We assume that the prices of the considered assets evolve according to the dynamics:

$$\begin{cases} dS_t = S_t\sigma(t, X_t)\, dW_t^1 \\ dX_t = -\mu(t, X_t)\, dt + \zeta(t, X_t)\, dW_t^1 + \xi(t, X_t)\, dW_t^2. \end{cases} \tag{2.1}$$

Here, uncertainty is due to a two dimensional standard Brownian motion (W_t) defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote $\{\mathcal{F}_t, 0 \leq t \leq T\}$ the usual augmented filtration of $\{W_t, 0 \leq t \leq T\}$. X_t can be viewed as a pure volatility asset, for instance a variance swap. In order to obtain a unique strong solution, one needs to assume some properties of the functions σ, μ and ξ.

$$\sigma^2, \mu, \zeta, \xi \text{ are locally Lipschitz and of linear growth on } (0, T) \times (0, +\infty) \tag{2.2}$$

Furthermore, we will need two other assumptions, in order to find the boundary condition of the pricing PDE:

$$\xi(., 0) = 0 \text{ and } \xi(., x) > 0 \text{ for } x \geq 0 \tag{2.3}$$

$$\zeta(., 0) = 0 \text{ and } \zeta(., x) > 0 \text{ for } x \geq 0 \tag{2.4}$$

$$\mu(t, x) \geq 0 \text{ on } [0, T] \times [0, +\infty) \tag{2.5}$$

$$\text{There exists a constant } C_\sigma \text{ such that } \sigma^2(t, x) \leq C_\sigma x \text{ for all}(t, x) \in [0, T) \times [0, +\infty) \tag{2.6}$$

These assumptions ensure that the process X remains nonnegative. Given those, one can prove that equation (2.1) has a unique strong solution $(S_u^{t,s,x}, X_u^{t,s,x})$ valued in $[0, +\infty)^2$ given:

$$S_t^{t,s,x} = s \ , \ X_t^{t,s,x} = x \ , \ (t, s, x) \in [0, T] \times [0, +\infty)^2$$

Remark 2.1. *It would be more realistic to consider the asset X as a call option, for example. But conditions (2.3) and (2.6) would have to be modified (by arbitrage, the value of a call option can not be below its discounted payoff). Modifying (2.3) would change the domain of the pricing PDE, in the case of a call option from $(0, T) \times (0, +\infty)^2$ to $(0, T) \times (0, +\infty) \times ((s - K)_+, s)$ Hence, the proof of the comparison principle would be more complicated. Changing condition (2.6) may change the limit condition of the value function near the boundary of the domain. It could make it more tedious to derive the equivalence of propositions 3.5 and 3.6. Hence we decide to study the simple case which* *embeds variance swaps, or futures on VIX index (futures contracts on the implied volatility level). Another issue is whether these results can be adapted to diffusion in dimension $n > 2$, with $k > 1$ constrained assets. If $k = 1$, there is not much work to adapt the case. Meanwhile, $k > 1$ would mean that each of the k boundary conditions would be the solution of the same kind of problem with $k - 1$ constrained assets, which would introduce some new difficulties.*

The aim of this paper is to derive a hedging price for a contingent claim $g(S_T)$ under certain constraints described in the following. For the sake of simplicity, we will consider some regularity assumptions on the payoff function:

$$g \text{ is bounded by a constant } C^* \tag{2.7}$$

$$g \text{ is } \mathcal{C}^2 \text{ and } s \to s^2 g"(s) \text{ is bounded by a constant } C_g \tag{2.8}$$

The second assumption could be relaxed with little efforts (considering a sequence of regular payoffs above the one of interest). Indeed, this assumption will only be used in the proofs of propositions 3.5 and 3.6, and one can see that for most common payoffs, these can be adapted.

2.2 The super-replication problem

Value function

The agent can trade assets on the market with self financing strategies, and its wealth process can be written as:

$$Y_r^{t,s,x,y,\pi} = y + \int_t^r \pi_S^{t,s,x}(u) dS_u^{t,s,x} + \int_t^r \pi_X^{t,s,x}(u) \left(dX_u^{t,s,x} + \mu(u, X_u^{t,s,x}) du \right)$$

Our problem is to find the super-replication price of a contingent claim $g(S_T)$ with a limited set of admissible strategies. One must find the minimum amount of money which enables to super-replicate de payoff of the option. Hence, the problem is to characterize the following value function:

$$v(t,s,x) = \inf_{y \in \mathbb{R}} \left\{ y : Y_T^{t,s,x,y,\pi} \geq g\left(S_T^{t,s,x}\right) \text{ a.s. for some } \pi \in \mathcal{A}_{t,s,x} \right\} \tag{2.9}$$

Gamma constraints

Here, we describe the set of admissible strategies $\mathcal{A}_{t,s,x}$. The specificity of our work is the following: we can buy and sell the asset S freely, without transaction cost or waiting time, but the asset X is far less liquid, so we need some time to by and sell it. Mathematically, this means that almost every adapted self-financed strategies (excepted doubling ones) will be admissible for the asset S, but that the set of admissible strategies will be far more constrained for the asset X. A trading strategy is a a vector $\pi(t) = (\pi_S(t), \pi_X(t))$, where $\pi_S(t)$ is the amount (in unity of assets) of assets S of the strategy at time t. π is in $\mathcal{A}_{t,s,x}$ if it is of the form:

$$\pi_S(r) = \sum_{n=0}^{N-1} y_s^n 1_{\tau_s^n > t} + \int_t^r \alpha_s(u) du + \int_t^r \gamma^{s,s}(u) dS_u^{t,s} + \int_t^r \gamma^{s,x}(u) dX_u^{t,x},$$

where τ_s^n are stopping times for each n, y_s^n are $\mathcal{F}_{\tau_s^n}$-measurable random variables, and $\alpha_s, \gamma_s^{s,s}$ and $\gamma_s^{s,x}$ are almost surely bounded adapted processes. While π_X satisfies:

$$\pi_X(r) = \sum_{n=0}^{N-1} y_x^n 1_{\tau_x^n > t} + \int_t^r \alpha_x(u) du,$$

with τ_s^n and α_s filling the same conditions as above. For technical reasons, $\gamma = (\gamma^{s,s}, \gamma^{s,x})$ above must be of the form:

$$\gamma = \sum_{n=0}^{N-1} z^n 1_{\tau_n \leq t < \tau n+1} + \int_t^r \psi_u^s du + \int_t^r \chi_u^s dS_u^{t,s} + \int_t^r \kappa_u^s dX_u^{t,x}$$

With ψ, χ, κ adapted and uniformly bounded. This is necessary to apply the result on double stochastic integrals proved in the section 6.

In other words, the terms γ with respect to X are constrained to be equal to zero. There are two main reasons to study these kind of constraints on π_X. As X is likely to be an illiquid asset, it introduces transaction costs. Therefore, any portfolio strategy such that $\gamma^{x,s} \neq 0$ or $\gamma^{x,x} \neq 0$ introduces infinite variation of the quantity of asset X held in the hedging portfolio. Therefore it would introduce infinite transaction costs, which are not acceptable. Hence, v can be viewed as a minorant of the super-replication price with vanishing transaction costs. On the other hand, as we will see in the following, these constraints induce robustness with respect to ξ and ζ as a byproduct. That is, constraining $(\gamma^{x,s}, \gamma^{x,x})$ allows the super-replication pricing and hedging to work even with misspecified ξ and ζ. This might be very useful, as ξ and ζ may be stochastic and driven by a factor against which one cannot hedge with the available assets.

2.3 Main results

Operators

First, we use the following notation:

$$\Sigma(t,s,x) = \begin{pmatrix} s\sigma(t,x) & 0 \\ 0 & \xi(t,x) \end{pmatrix}$$

Remark that we did not include parameter ζ on purpose. The operator used to to define the super-replication price equation will be:

$$\begin{aligned} & F\left(t,s,x,Du,D^2u\right) = \\ \lambda^- & \begin{pmatrix} -\frac{\partial u}{\partial t} + \mu(t,x)\frac{\partial u}{\partial X} - \frac{1}{2}s^2\sigma^2(t,x)\frac{\partial^2 u}{\partial S^2} & -\frac{1}{2}s\sigma(t,x)\xi(t,x)\frac{\partial^2 u}{\partial S \partial X} \\ -\frac{1}{2}s\sigma(t,x)\xi(t,x)\frac{\partial^2 u}{\partial S\partial X} & -\frac{1}{2}\left(\xi(t,x)\right)^2\frac{\partial^2 u}{\partial X^2} \end{pmatrix} \\ & = \lambda^-\left(J\left(t,s,x,Du,D^2u\right)\right), \end{aligned} \tag{2.10}$$

where λ^- represents the smallest eigenvalue of a matrix. One can easily check that this operator is parabolic. One can derive a more intuitive formulation for this operator. Indeed, by elementary properties of symmetric matrices, one can see that $F\left(t,s,x,Du,D^2u\right)$ has the same sign as the expression:

$$-\frac{\partial u}{\partial t} + \mu(t,x)\frac{\partial u}{\partial X} - \frac{1}{2}s^2\sigma^2(t,x)\frac{\partial^2 u}{\partial S^2} - \sup_{\xi\in\mathbb{R}}\left\{s\sigma(t,x)\xi\frac{\partial^2 u}{\partial S\partial X} + \frac{1}{2}\xi^2\frac{\partial^2 u}{\partial X^2}\right\}.$$

This expression reminds regular control problems, but it is not usable here as it may take infinite values. In the following, we will denote the matrix H^ξ as:

$$H^\xi\left(t,s,x,Du,D^2u\right) = \begin{pmatrix} -\frac{\partial u}{\partial t} + \mu(t,x)\frac{\partial u}{\partial X} - \frac{1}{2}s^2\sigma^2(t,x)\frac{\partial^2 u}{\partial S^2} & -\frac{1}{2}S\sigma(t,x)\xi\frac{\partial^2 u}{\partial S\partial X} \\ -\frac{1}{2}S\sigma(t,x)\xi\frac{\partial^2 u}{\partial S\partial X} & -\frac{1}{2}\xi^2\frac{\partial^2 u}{\partial X^2} \end{pmatrix}. \tag{2.11}$$

By observing that a 2×2 matrix is positive if and only if its diagonal terms and its determinant are positive, we see that operator F is positive if and only if matrix H^ξ is positive for a given $\xi \neq 0$. Furthermore, we see that ζ does not enter into account in these operators.

Equations

Now, we can state the two main results of this paper:

Proposition 2.1. *The solution v of the super-replication problem is the unique viscosity solution of the equation:*

$$F\left(t,s,x,Dv,D^2v\right) = 0 \text{ on } (0,T)\times(0,+\infty)^2$$

such that v is continuous on the boundaries $x=0$ and $t=T$, with $v(t,s,0) = v(T,s,x) = g(s)$ for all $(t,s,x)\in[0,T]\times[0,+\infty)^2$.

The second main result is a dual representation theorem of the value function.

Proposition 2.2. *The solution v of the super-replication problem satisfies the dual representation:*

$$v(t,s,x) = \sup_{(\rho,\xi)\in\mathcal{U}} \mathbf{E}\left[g\left(S_T^{t,s,x,\rho,\xi}\right)\right] \text{ for all } (t,s,x)\in[0,T]\times[0,+\infty)^2$$

where $S_\cdot^{t,s,x,\rho,\xi}$ is the solution of the equation:

$$\begin{aligned} dS_u^{t,s,x,\rho,\xi} &= \sigma(t,X_u^{t,x,\rho,\xi})S_u^{t,s,x,\rho,\xi}dW_u^1 \\ dX_u^{t,x,\rho,\xi} &= -\mu(t,X_u^{t,x,\rho,\xi})du + \xi(u)X_u^{t,x,\rho,\xi}dW_u^2 \\ \left\langle dW_u^1, dW_u^2\right\rangle &= \rho(u) \\ S(t) = s\ &,\ X(t) = x \end{aligned}$$

and $\mathcal{U}$ is the set of all almost-surely bounded progressively measurable processes taking values in $[-1;1]\times[0;+\infty)$

3 Viscosity property

3.1 Sub and supersolution characterization

The proof of the viscosity property is very close to the proof in [23]. Though there are two noticeable differences. The first one is that, here, the space of gamma constraints is of empty interior, because the "gamma" with respect to the second asset is constrained to be zero. It has some impact on the proof of the subsolution property, but most of all on the uniqueness theorem. The second difference is that we did not suppose that the matrix of the gammas was symmetric (i.e. we did not suppose $\gamma_u^2 = 0$, while we constrained the gamma component of X to be equal to zero), so we will need to study the small time behavior of double stochastic integrals involving non-symmetric matrices, which is a new feature. The proof of the sub and supersolution properties involves respectively two auxiliary value functions $\overline{v} \geq v$ and $\underline{v} \leq v$, and are quite long and technical[1]. Furthermore, the characterization of v is found by the comparison theorem, that gives $\overline{v} \leq \underline{v}$ leading

$$\overline{v} = v = \underline{v}$$

But here, for the sake of simplicity we will only give the main arguments of the proof, without involving rigorous mathematics. But the same steps as in [23] could be used. We will act as if we manipulated the original value function. One could object that we prove the boundary properties in the next sections for v and that we should do it for $\overline{v}$ and $\underline{v}$. But the proof would be exactly the same, as the differences between the definitions of the three value functions would not interfere.

Definition 3.1. *Let w be a locally bounded function. Let w_* (resp w^*) be its lower (resp upper) semicontinuous envellopes.*
w is a viscosity supersolution of (2.10) if, for any $(t_0, s_0, x_0) \in [0,T) \times (0,+\infty)^2$:

$$F\left(t_0, y_0, D\varphi(t_0, y_0), D^2\varphi(t_0, y_0)\right) \geq 0 \tag{3.1}$$

for all $\varphi \in C^\infty\left([0,T) \times (0,+\infty)^2\right)$ such that

$$0 = (w_* - \varphi)(t_0, y_0) = \min_{(t,y)[O,T[\times[0,+\infty[^2} (w_* - \varphi)(t, y)$$

And w a viscosity subsolution of (2.10) if, for any $(t_0, s_0, x_0) \in [0,T) \times (0,+\infty)^2$:

$$F\left(t_0, y_0, D\varphi(t_0, y_0), D^2\varphi(t_0, y_0)\right) \leq 0 \tag{3.2}$$

for all $\varphi \in C^\infty\left([0,T) \times (0,+\infty)^2\right)$ such that

$$0 = (w^* - \varphi)(t_0, y_0) = \max_{(t,y)[O,T[\times[0,+\infty[^2} (w`* - \varphi)(t, y)$$

Subsolution property

Let us begin by defining the upper bound $\overline{v}$ for the value function v. First, we define a norm on the controls:

$$\|\nu\|_{t,s}^{\beta,\infty} := \max\left\{\|N\|_{L^\infty} ; \|Y\|_{t,s}^{\beta,\infty} ; \|\alpha\|_{t,s}^{\beta,\infty} ; \|\gamma\|_{t,s}^{\beta,\infty} ; \|\gamma^{s,s}\|_{t,s}^{\beta,\infty} ; \|\gamma^{s,s}\|_{t,s}^{\beta,\infty}\right\}.$$

We define another set of admissible controls with:

$$\mathcal{A}_{t,s,x}^M = \left\{\nu \in \mathcal{A}_{t,s,x} : \|\nu\|_{t,s}^{\beta,\infty} \leq M\right\},$$

and the auxiliary value function:

$$v^M(t,s) := \inf\left\{y \in \mathbb{R} : X_{t,s,x,y}^v(T) \geq g(S_{t,s,x}(T)) \text{ for some } \nu \in \mathcal{A}_{t,s,x}^M\right\}.$$

[1]I would like to thank Nizar Touzi for his explanations about this topic

We get an upper bound of v by taking:

$$\overline{v} := \inf_{M>0} (v^M)^*(t, s, x).$$

Now we can state the viscosity subsolution property:

Proposition 3.3. *The function $\overline{v}$ is a viscosity subsolution of equation*

$$F(t, s, x, D\overline{v}, D^2\overline{v}) = 0 \text{ on } (0, T) \times (0, +\infty)^2.$$

Proof. The idea of the proof is that, if the function is not a viscosity solution, we can exhibit a strategy of super replication that costs less than the value function. Thus it leads to a contradiction. This is an adaptation of the proof in [23]. We omit some technical condition, which can be transcribed easily. There are some differences, thought, as the space of controls is of empty interior. As the payoff function is bounded, we know that the value function is finite. Let $\varphi \in C^\infty$ be a test function such that:

$$0 = (\overline{v} - \varphi)(t_0, s_0, x_0) > (\overline{v} - \varphi)(t, s, x) \text{ for all } (t, s, x) \neq (t_0, s_0, x_0).$$

Then assume that on the contrary

$$F(t_0, s_0, x_0, D\varphi, D^2\varphi) > 0. \tag{3.3}$$

We will obtain a contradiction. Denote:

$$\sigma(t_0, x_0) = \sigma_0 \ , \ \mu(t_0, x_0) = \mu_0.$$

First, remark that (3.3) leads to:

$$\frac{1}{2} s_0^2 \sigma_0^2 \frac{\partial^2 \varphi}{\partial S^2}(t_0, s_0, x_0) - \mu_0 \frac{\partial \varphi}{\partial X}(t_0, s_0, x_0) + \frac{\partial \varphi}{\partial t}(t_0, s_0, x_0) < 0. \tag{3.4}$$

For $\varepsilon > 0$ sufficiently small, we consider a compact neighborhood $\mathcal{N}$ of (t_0, s_0, x_0) such that:

$$\begin{aligned} &- \frac{\sigma^2(t,x)s^2}{\sigma_0^2 s_0^2} \left(\frac{\partial \varphi}{\partial t}(t_0, s_0, s_0) - \mu_0 \frac{\partial \varphi}{\partial X}(t_0, s_0, x_0) + \varepsilon \right) \\ &- \frac{\partial \varphi}{\partial X}(t_0, s_0, s_0)\mu(t, x) + \max_{s,x \in \mathcal{N}} \frac{\partial \varphi}{\partial t}(t, s, x) \leq 0 \end{aligned} \tag{3.5}$$

$$\begin{pmatrix} \frac{1}{2}\sigma_0^2 s_0^2 \frac{\partial^2 \varphi}{\partial S^2}(t,s,x) + \left(\frac{\partial \varphi}{\partial t} - \mu_0 \frac{\partial \varphi}{\partial X} \right)(t_0, s_0, x_0) + \varepsilon & \frac{1}{2} s_0 \sigma_0 \xi_0 \frac{\partial^2 \varphi}{\partial S \partial X}(t,s,x) \\ \frac{1}{2} s_0 \sigma_0 \xi_0 \frac{\partial^2 \varphi}{\partial S \partial X}(t,s,x) & \frac{1}{2}\xi_0^2 \frac{\partial^2 \varphi}{\partial X^2}(t,s,x) \end{pmatrix} \leq 0 \tag{3.6}$$

for all $(t, s, x) \in \mathcal{N}$.

As φ is C^∞ and satisfies (3.3), $\mathcal{N}$ is nonempty and $(t_0, s_0, x_0) \notin \partial\mathcal{N}$ for sufficiently small ε. As (t_0, s_0, x_0) is a strict maximizer of $\overline{v} - \varphi$, there exists $\eta > 0$ such that $(\overline{v} - \phi)(t, s, x) < 2\eta$ on $\partial\mathcal{N}$. Let θ be the stopping time:

$$\theta := \inf \{t \geq t_0 : (t, S_t, X_t) \notin \mathcal{N}\},$$

and consider the following decomposition of φ into a super replicable part and a negative part:

$$
\begin{aligned}
\varphi(t,s,x) &= \psi_1(t,s,x) + \psi_2(t,s,x) \\
\psi_1(t,s,x) &= \varphi(t_0,s_0,x_0) - \frac{(s-s_0)^2}{s_0^2\sigma_0^2}\left(\frac{\partial\varphi}{\partial t}(t_0,s_0,x_0) + \varepsilon - \mu_0\frac{\partial\varphi}{\partial X}(t_0,s_0,x_0)\right) \\
&\quad + (x-x_0)\frac{\partial\varphi}{\partial X}(t_0,s_0,x_0) + (s-s_0)\frac{\partial\varphi}{\partial S}(t_0,s_0,x_0) \\
&\quad + \int_{t_0}^{t} \max_{(u',s',x')\in\mathcal{N}} \frac{\partial\varphi}{\partial t}(u',s',x')du \\
\psi_2(t,s,x) &= \varphi(t,s,x) - \varphi(t_0,s_0,x_0) \\
&\quad + \frac{(s-s_0)^2}{s_0^2\sigma_0^2}\left(\frac{\partial\varphi}{\partial t}(t_0,s_0,x_0) + \varepsilon - \mu_0\frac{\partial\varphi}{\partial X}(t_0,s_0,x_0)\right) \\
&\quad - (s-s_0)\frac{\partial\varphi}{\partial S}(t_0,s_0,x_0) - (x-x_0)\frac{\partial\varphi}{\partial X}(t_0,s_0,x_0) \\
&\quad - \int_{t_0}^{t} \max_{(u',s',x')\in\mathcal{N}} \frac{\partial\varphi}{\partial t}(u',s',x')du.
\end{aligned}
$$

- First, let us prove that one can super-replicate the first part of this decomposition. Consider the initial capital:

  $$y_0 := \overline{v}(t_0,s_0,x_0) - \eta,$$

  and the control:

  $$\pi_0 = D\varphi(t_0,s_0,x_0), \alpha(t) := 0,$$
  $$\pi_S(t) = -\frac{2(S-S_0)}{S_0^2\sigma_0^2}\left[\left(\frac{\partial\varphi}{\partial t} - \mu_0\frac{\partial\varphi}{\partial X}\right)(t_0,s_0,x_0) + \varepsilon\right].$$

  Denote the portfolio strategy $(Y,\pi) := (Y^{\nu}_{t_0,s_0,x_0,y_0}, \pi^{\nu}_{t_0,s_0,x_0})$. Then, by Ito's formula, combined with conditions (3.4) and (3.5), we have:

  $$
  \begin{aligned}
  dY(t) &= \frac{\partial\varphi}{\partial X}(t_0,s_0,x_0)\,(dX + \mu(t,X)dt) \\
  &\quad - \frac{2(S-s_0)}{s_0^2\sigma_0^2}\Big[\Big(\frac{\partial\varphi}{\partial t} - \mu_0\frac{\partial\varphi}{\partial X}\Big)(t_0,s_0,x_0) + \varepsilon\Big]dS \\
  d\Psi_1(t,S_t,X_t) &= -\frac{2(S-s_0)}{s_0^2\sigma_0^2}\left[\left(\frac{\partial\varphi}{\partial t} - \mu_0\frac{\partial\varphi}{\partial X}\right)(t_0,s_0,x_0) + \varepsilon\right]dS + \max_{\mathcal{N}}\frac{\partial\varphi}{\partial t}dt \\
  &\quad + \frac{2S^2\sigma^2(t,X)}{S_0^2\sigma_0^2}\left[\left(\frac{\partial\varphi}{\partial t} - \mu_0\frac{\partial\varphi}{\partial X}\right)(t_0,s_0,x_0) + \varepsilon\right]dt \\
  & d(Y(t) - \psi_1(t,S(t),X(t))) \geq 0.
  \end{aligned}
  $$

  This shows that $Y(\theta) - \psi_1(\theta,S(\theta),X(\theta)) \geq -\eta$.

- Now, let us show that $\psi_2 \leq 0$ on $\mathcal{N}$. For $t = t_0$, we have by differentiating with respect to s and x:

  $$D\psi_2\,(t_0,s_0,x_0) = 0.$$

The Hessian matrix of ψ_2 is given by: :

$$D^2\psi_2(t,s,x) = \begin{pmatrix} \frac{\partial^2\varphi}{\partial S^2}(t,s,x) + \frac{2}{s_0^2\sigma_0}\left(\left(\frac{\partial\varphi}{\partial t} - \mu_0\frac{\partial\varphi}{\partial X}\right)(t_0,s_0,x_0)+\varepsilon\right) & \frac{\partial^2\varphi}{\partial S\partial X}(t,s,x) \\ \frac{\partial^2\varphi}{\partial S\partial X}(t,s,x) & \frac{\partial^2\varphi}{\partial x^2}(t,s,x) \end{pmatrix},$$

for any $(t_0,s,x)\in\mathcal{N}$. Let $\Sigma_0 = \begin{pmatrix} \sigma_0 s_0 & 0 \\ 0 & \xi_0 \end{pmatrix}$. From assumption (3.5), we get:

$$\frac{1}{2}\Sigma_0 D^2\psi_2(t_0,s_0,x_0)\Sigma_0 = \begin{pmatrix} \frac{1}{2}\sigma_0^2 s_0^2\frac{\partial^2\varphi}{\partial S^2}(t,s,x) + \left(\frac{\partial\varphi}{\partial t} - \mu_0\frac{\partial\varphi}{\partial X}\right)(t_0,s_0,x_0)+\varepsilon & \frac{1}{2}s_0\sigma_0\xi_0\frac{\partial^2\varphi}{\partial S\partial X}(t,s,x) \\ \frac{1}{2}s_0\sigma_0\xi_0\frac{\partial^2\varphi}{\partial S\partial X}(t,s,x) & \frac{1}{2}\xi_0^2\frac{\partial^2\varphi}{\partial X^2}(t,s,x) \end{pmatrix} \leq 0$$

We deduce that the function ψ_2 is concave on $\mathcal{N}\cap\{t=t_0\}$, and its value and first order derivative are 0 at (t_0,s_0,x_0). Hence, it is negative on $\mathcal{N}$ for $t=t_0$. Now, remark that its time derivative is negative on $\mathcal{N}$, and so, ψ_2 is negative on $\mathcal{N}$ for $t>t_0$.

Therefore, we have:

$$Y(\theta) \geq \varphi(\theta,S_\theta,X_\theta) - \eta \geq v(\theta,S_\theta,X_\theta),$$

and the dynamic programming principle is violated. This concludes the proof. □

Supersolution property

To prove the supersolution property, one has to define the relaxed stochastic control problem, for any $M>0$:

$$\underline{v}^M(t,s,x) := \inf\Big\{y\in\mathbb{R} : \tilde{Y}^{\nu}_{t,s,x,y}(T) \geq g(\tilde{S}_{t,s,x}(T)) \text{ for some } (\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{F}},\tilde{P}) \text{ and } \tilde{\nu}\in\mathcal{A}^M_{t,s}(\tilde{\Omega})\Big\}.$$

Here, $\tilde{Y},\tilde{S}$ and $\mathcal{A}^M_{t,s}(\tilde{\Omega})$ are defined as in the original control problem. Then, define $\underline{v}$ as the lower semicontinuous envelope of the inferior bound of $\underline{v}^M$ over all $M>0$. The change of probability and filtration is due to technical reason in order to obtain existence of and optimal control, and lower semicontinuity of the value function. These, and the corresponding dynamic programming principle are obtained in lemmas 5.1 and 5.2 in [23].

Proposition 3.4. *For all M sufficiently large, $\underline{v}^M$ is a viscosity supersolution of equation:*

$$F(t,s,x,D\underline{v}^M,D^2\underline{v}^M) = 0 \text{ on } [0,T)\times(0,+\infty)^2. \tag{3.7}$$

Proof. This proof is exactly like in Theorem 5.4 in [23], excepted for the limit result on double stochastic integrals. Indeed, in that paper, the integrand of the double integral is supposed to be symmetric, whereas here it is not. But anyway the result is the same, as the integrand turns out to be necessarily symmetric. For these reasons, we only give a sketch of the demonstration. Let $M \geq C^*$ (where C^* is the bound of g in assumption (2.7)) be fixed. By lemma 5.2, in [23] $\underline{v}^M$ is finite and lower semicontinious. Consider a $(t_0, s_0, x_0) \in [0,T) \times \mathbb{R}_+^2$ and a test function $\varphi \in C^\infty[0,T] \times \mathbb{R}_+^2$ such that:

$$0 = (\underline{v}^M - \varphi)(t_0, s_0, x_0) = \min_{(t,s,x)\in[0,T[\times\mathbb{R}_+^2} (\underline{v}^M - \varphi)(t, s, x)$$

Set $y_0 = \underline{v}^M(t_0, s_0, x_0)$. By Lemma 5.2 in [23], there exists a two dimensional brownian motion $\tilde{W}$ on a filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{P})$ satisfying the usual conditions and a control $\pi \in \mathcal{A}^M_{t_0,s_0,x_0}$ such that, for any stopping time $t_0 \leq \theta \leq T$:

$$\tilde{Y}^{\tilde{\pi}}_{t_0,s_0,x_0}(\theta) \geq \underline{v}^M(\theta, \tilde{S}(\theta), \tilde{X}(\theta)) \geq \varphi(\theta, \tilde{S}(\theta), \tilde{X}(\theta)).$$

By twice applying Ito's lemma, one obtains, by denoting $\tilde{Z} = \begin{pmatrix} \tilde{S} \\ \tilde{X} \end{pmatrix}$:

$$\int_{t_0}^{\theta} l(r)dr + \int_{t_0}^{\theta} \left(c + \int_{t_0}^{r} a(u)du + \int_{t_0}^{r} b(u)d\tilde{Z}(u) \right)^T d\tilde{Z}(r) \geq 0, \tag{3.8}$$

where:

$$\begin{aligned} l(r) &:= -\mathcal{L}\varphi \left(r, \tilde{S}(r), \tilde{X}(r) \right) \\ a(r) &:= \tilde{\alpha}(r) - \mathcal{L}(D\varphi) \left(r, \tilde{S}(r), \tilde{X}(r) \right) \\ b(r) &:= \tilde{\gamma}(r) - \mathcal{L}(D^2\varphi) \left(r, \tilde{S}(r), \tilde{X}(r) \right) \\ c &:= \tilde{\pi}^0 - D\varphi(t_0, s_0, x_0). \end{aligned}$$

In [23] it is then proved that $c = 0$ and by considering $\theta^\eta = \min(\theta, \eta)$, we have for any real number $\varepsilon > 0$:

$$\lim_{\eta \to 0^+} \eta^{\varepsilon - 3/2} \int_{t_0}^{\theta^\eta} \left(\int_{t_0}^{r} a(u)du \right)^T d\tilde{Z}(r) = 0. \tag{3.9}$$

Then it follows from (3.8) that:

$$\liminf_{\eta \to 0^+} \frac{1}{\eta \log\log \frac{1}{\eta}} \int_{t_0}^{\theta^\eta} \left(\int_{t_0}^{r} b(u)d\tilde{Z}(u) \right)^T d\tilde{Z}(r) \geq 0. \tag{3.10}$$

Therefore, it follows from proposition 6.14 that:

$$b(t_0) \text{ is symmetric and positive.} \tag{3.11}$$

Hence, by definition of b, $\gamma_0^{s,x} = 0$. Denoting

$$\Sigma_0 = \begin{pmatrix} s_0\sigma(t_0, x_0) & 0 \\ 0 & \xi(t, x_0) \end{pmatrix} \tag{3.12}$$

we get that the following matrix is positive:

$$\Sigma_0 b(t_0) \Sigma_0 = \begin{pmatrix} -\sigma_0^2 s_0^2 \left(\frac{\partial^2\varphi}{\partial S^2} - \gamma_0^{s,s} \right) & -\sigma_0 s_0 \xi_0 \frac{\partial^2\varphi}{\partial S \partial X} \\ -\sigma_0 s_0 \xi_0 \frac{\partial^2\varphi}{\partial S \partial X} & -\xi_0^2 \frac{\partial^2\varphi}{\partial X^0} \end{pmatrix} \geq 0, \tag{3.13}$$

where all the derivatives are taken at point (t_0, s_0, x_0). Finally, theorem A.2 in [23] shows:

$$\limsup_{\eta\to 0^+} \frac{1}{\eta} \int_{t_0}^{\theta^\eta} \left(\int_{t_0}^{r} b(u) d\tilde{Z}(u) \right)^T d\tilde{Z}(r) = \frac{1}{2} Tr \left[\Sigma_0 b(t_0) \Sigma_0 \right]$$

Dividing (3.8) by η, recalling (3.9) and that $c = 0$, and taking the limit $\eta \to 0$ we obtain:

$$-\mathcal{L}\varphi(t_0, s_0) - \frac{1}{2} Tr \left[\Sigma_0 b(t_0) \Sigma_0 \right] \geq 0,$$

and so:

$$-\frac{\partial \varphi}{\partial t} + \mu(t_0, x_0) \frac{\partial \varphi}{\partial X} - \frac{1}{2} \sigma^2(t_0, x_0) s_0^2 \gamma_0^{s,s} \geq 0$$

Plugging it into (3.13) finishes the proof. □

Then, one can prove that $\underline{v}$ is supersolution of the same equation, using the same steps as in Corollary 5.5 in [23].

3.2 Boundary conditions

Following remark of example 4.1 in the last section, the viscosity property of the value function v in the interior of the domain is not enough to ensure the characterization of v. Indeed, one needs to derive the boundary behavior near the boundary $x = 0$ to obtain uniqueness of the solution of equation (2.10) with this additional constraint. This is why we need assumptions (2.3) and (2.5) to exhibit a superhedging strategy when $X \to 0$, which gives an upper bound on v.

Terminal condition

In many super-replication problems, the value function converges to a face-lifted payoff when time tends to maturity. This is not the case here. Let us demonstrate that the terminal condition of v corresponds to the payoff function g.

Proposition 3.5. *The terminal condition of the value function v is g. In other words: For any $(s, x) \in \mathbb{R}_+^2$*

$$\lim_{t \nearrow T, s' \to s, x' \to x} v(t, s', x') = g(s) \tag{3.14}$$

That is, the value function is continuous on $t = T$.

Proof. Recall assumptions (2.5),(2.8) and (2.6). Consider an instant $t < T$ and a state of the market (s, x). Consider the following portfolio for any time $t \leq u \leq T$, which will be the key of the demonstrations below:

$$\pi_u = \begin{pmatrix} \pi_u^S \\ \pi_u^X \end{pmatrix} = \begin{pmatrix} g'(S_u) \\ \frac{1}{2} C_\sigma C_g (T - u) \end{pmatrix}. \tag{3.15}$$

Here, C_g is the constant in assumption (2.8). On the other hand, by Ito's formula one has:

$$g(S_T) - g(S_t) = \int_t^T g'(S_u) dS_u + \frac{1}{2} \sigma^2(u, X_u) S_u^2 g''(S_u) du.$$

Plugging conditions (2.6) and (2.8) one obtains:

$$g(S_T) - g(S_t) \leq \int_t^T g'(S_u)dS_u + \frac{1}{2}C_\sigma C_g X_u du. \tag{3.16}$$

Starting with the initial wealth $g(S_t) + \frac{1}{2}X_t C_\sigma C_g(T-t)$, the continuous selling of X_u gives:

$$\int_t^T \frac{1}{2}C_\sigma C_g X_u du,$$

while, with condition (2.5) we obtain positive dividends as $\mu \geq 0$. Therefore the profit and loss associated with the component in X (excluded the buying price at the beginning) of the portfolio dominates:

$$\int_t^T \frac{1}{2}C_\sigma C_g X_u du.$$

By combining with (3.16), one gets that wealth $g(S_t) + \frac{1}{2}xC_\sigma C_g(T-t)$ is enough to super-replicate the payoff. Hence:

$$v(t,s,x) \leq g(s) + \frac{1}{2}xC_\sigma C_g(T-t).$$

The reverse inequality is more usual, and comes from the fact that v is dominated by the replication price $u(t,s,x)$ without constraints, which is the expectation of the payoff. We will not prove this assertion here as it is classical. Once this is done, applying Fatou's lemma finishes the proof:

$$u(t,s,x) \leq v(t,s,x) \leq g(s) + \frac{1}{2}xC_\sigma C_g(T-t).$$

Then as the LHS and the RHS converge to g, the value function does too. □

Lateral condition

The next proposition deals with the same type of conditions near $x = 0$.

Proposition 3.6. *The boundary condition of the value function v near $x = 0$ is g: For any $(s,t) \in \mathbb{R}_+^2$*

$$\lim_{t' \to t, s' \to s, x' \to 0} v(t', s', x') = g(s). \tag{3.17}$$

Proof. The proof is essentially the same as above. □

4 The comparison result

In this section, we prove that equation (2.10) has a unique solution, by establishing a comparison result. Our proof mostly relies on a strict supersolution argument, which has been introduced by Ishii and Lions in [48] and used by Soner et al. in [23]. The idea is to prove a comparison for perturbed sub and super-solutions, and then to take the limit of the resulting inequalities when the perturbation tends to zero. But first, we will see under which conditions does the comparison principle hold.

4.1 Boundary conditions

Interestingly, unlike in most similar parabolic problems, one will not only need a terminal condition to obtain uniqueness, but also some boundary conditions, when the spot price and the volatility asset tend to zero. Another condition is naturally introduced by the fact that we only consider bounded solutions. This is because equation (2.10) is not parabolic in the most common sense, due to a nonlinearity in front of the time derivative. Here is a simple example to illustrate this fact:

Example 4.1. *This equation, defined for $u(t,s)$, $(t,s) \in [O,T] \times [0,+\infty[$:*

$$min \left\{ \frac{\partial v}{\partial t} + S^2 \frac{\partial^2 v}{\partial S^2}, \frac{\partial^2 v}{\partial S^2} \right\} = 0 \tag{4.1}$$

$$u(T) = 0 \tag{4.2}$$

has no unique solution. Indeed, let us consider two families of functions:

$$u(t,s) = (t-T)\lambda \text{ and } u(t,s) = (t-T)s\lambda$$

With $\lambda \geq 0$. Both are solutions of equation (4.1). In order to eliminate these solution we need to impose a condition like:

$$u(t,0) = 0 \text{ for all } t \in [0,T]$$

To eliminate the first kind of solution, and a more common condition

$$u \text{ is bounded on } [0,T] \times [0,+\infty[$$

For the second one. Then, with these boundary conditions, using the following method, one can prove that $u = 0$ is now the only solution of equation (4.1) in the viscosity sense.

This is why one must use boundary conditions (3.14) and (3.17)

$$\lim_{t \to T^-} v(t,s,.) = g(s) \ , \ \lim_{X \to 0^+} v(.,s,x) = g(S) \ ,$$

and v is bounded by a constant C.

4.2 Equivalent equation

In order to establish the comparison result, we can reformulate the operator (2.10) with $\xi(t,x) = \max(1,x)$. We can easily see that changing the operator leaves the equation unchanged on the open domain $(0,+\infty)^2$, because of assumption (2.3). Indeed, changing $\xi(t,X) > 0$ into another positive function does not change the sign of the operator F for fixed $\left(t,s,x,Du,D^2u\right)$, with $x > 0$. So we can introduce a new assumption to prove the uniqueness theorem:

$$\xi(t,x) = \max(1,x) \text{ for all } x \in (0,+\infty). \tag{4.3}$$

Note that this is only a notation to rewrite the PDE in an equivalent way. It is not meant to describe the dynamics of the process X.

4.3 Strict viscosity supersolutions

Let us now introduce the notion of strict supersolution, as in [23] and [48]. This strict supersolution property will be necessary to prove the comparison principle.

Definition 4.2. *For a strictly positive constant η, a function w is an η-strict viscosity supersolution of equation (2.10) if:*

$$F\left(t_0, y_0, D\varphi(t_0,y_0), D^2\varphi(t_0,y_0)\right) > \eta \tag{4.4}$$

for all (t_0, y_0) $[0,T) \times (0,+\infty)^2$ and $\varphi \in C^\infty\left([0,T) \times (0,+\infty)^2\right)$ such that

$$0 = (w_* - \varphi)(t_0, y_0) = \min_{(t,y)[O,T[\times[0,+\infty[^2} (w_* - \varphi)(t,y)$$

The next step is to find a function $w^1 \geq 0$ which one can add to any viscosity supersolution w of (2.10) to build a superior and arbitrary close strict supersolution $w + \varepsilon w^1$. As we will prove a comparison result for strict supersolutions, the next lemma will enable us to manage the comparison with any non-strict supersolution. Indeed, by perturbing the supersolution and taking the limit when the perturbation tends to zero, one can extend comparison. The main difficulty is that w^1 must always be superior to zero, and be concave enough to have $\frac{\partial^2 w^1}{\partial X^2}$ sufficiently negative to ensure property (4.4).

Lemma 4.1. *Assume (4.3). Then the function*

$$w^1(t,s,x) := (T-t) + \ln(1+x) \geq 0$$

Is a η-strict viscosity supersolution of (2.10) on $[0,T] \times (0,+\infty)^2$ for some $\eta > 0$. Furthermore, if w is a supersolution of (2.10) with $w(T,.) \geq g$, then, for any $\epsilon > 0$, $w + \epsilon w^1$ is a $\epsilon^2\eta$-strict supersolution of (2.10) with $(w+\epsilon w^1)(T,.) \geq g$.

Proof. One can easily check that w^1 is a strict supersolution:

$$\begin{aligned} F\left(t,s,x,Dw^1,D^2w^1\right) &= \lambda^- \begin{pmatrix} 1 + \frac{\mu(x)}{1+x} & 0 \\ 0 & \max\left(1,x^2\right)\frac{1}{(1+x)^2} \end{pmatrix} \\ &\geq 1 \end{aligned}$$

Now, we check that $w + \varepsilon w^1$ is a ε-strict supersolution. Indeed, for any $(t_0, s_0, x_0) \in (0,T) \times (0,+\infty)^2$, and for any test function $\varphi :\in C^\infty\left((0,T) \times (0,+\infty)^2 \to \mathbb{R}\right)$ which satisfies:

$$Min\left(\varphi - w - \varepsilon w^1\right) = \left(\varphi - w - \varepsilon w^1\right)(t_0, s_0, x_0) = 0$$

Then, as $w^1 \in C^\infty\left((0,T) \times (0,+\infty)^2 \to \mathbb{R}\right)$, $\psi = \varphi - \varepsilon w^1$ is a test function for w such that $\psi - w$ attains its strict minimum at (t_0, s_0, x_0). Hence, we have, by the supersolution property of w:

$$F\left(t_0, s_0, x_0, D\psi, D^2\psi\right) \geq 0$$

Then, considering that for any symmetric matrices: $\lambda^-(A+B) \geq \lambda^-(A) + \lambda^-(B)$

$$\begin{aligned} F\left(t_0, s_0, x_0, D\varphi, D^2\varphi\right) &\geq F\left(t_0, s_0, x_0, D\psi, D^2\psi\right) + F\left(t_0, s_0, x_0, D\left(\varepsilon w^1\right), D^2\left(\varepsilon w^1\right)\right) \\ &\geq F\left(t_0, s_0, x_0, D\psi, D^2\psi\right) + \varepsilon F\left(t_0, s_0, x_0, Dw^1, D^2w^1\right) \\ &\geq \varepsilon \end{aligned}$$

By homogeneity of F in w. □

Before approaching the technical proof of the comparison principle involving strict-supersolutions, let us see how the preceding lemma allows us to extend that principle to any supersolution, thus proving the main theorem that follows:

Proposition 4.7. *If w and u are respectively super and subsolution of (2.10), and there exists a function h such that for any $(t,s) \in [0,T] \times \mathbb{R}_+$:*

$$\lim \sup_{(t',s',x') \to (t,s,0)} u \leq h(t,s) \leq \lim \inf_{(t',s',x') \to (t,s,0)} w$$

and a function g such that for any s in$\mathbb{R}_+$:

$$\lim \sup_{(t',s',x') \to (T,s,x)} u \leq g(t,x) \leq \lim \inf_{(t',s',x') \to (T,s,x)} w$$

Then $u^* \leq w_*$ on $[0,T] \times (0,+\infty)^2$. In particular, the solution of equation (2.10) in the viscosity sense with boundary conditions is unique.

Proof. We use the same technique as in [23]. If w and u are respectively super and subsolutions of (2.10). Furthermore suppose that they both verify the limit conditions (3.14) and (3.17). Then, for any $\varepsilon > 0$, with lemma 4.1, $w + \varepsilon w^1$ satisfy the boundedness, strict supersolution and boundary limits assumptions of theorem 4.8. Applying it, one gets:

$$w + \varepsilon w^1 \geq u \text{ on } (0,+\infty)^2 \times [0;T]$$

Finally, letting ε converge to zero by positive values, we get the result:

$$w \geq u \text{ on } (0,+\infty)^2 \times [0;T]$$

□

4.4 Modulus of continuity of F

We now introduce some technical lemmas which are classical in the viscosity solutions theory. We need a modulus of continuity for the operator F. It is given in the next lemma:

Lemma 4.2. *Let A and $A' \in S^2(\mathbb{R})$ such that:*

$$-3\alpha \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} A & 0 \\ 0 & -A' \end{pmatrix} \leq 3\alpha \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} \tag{4.5}$$

And consider the function:

$$f_1^\varepsilon(s,s',x,x') = (s+s'+x+x')\varepsilon - (\ln(s)+\ln(s'))\varepsilon^2$$

For any (t,s,x,p) and (t',s',x',p') for which:

- *There exists a constant C_1 such that:*

$$\sigma^2(x) \leq \frac{C_1}{2\varepsilon} \text{ and } \sigma^2(x') \leq \frac{C_1}{2\varepsilon}$$

- *There exists a constant C_2^K such that Σ and μ are Lipschitz of constant C_2^K on a convex domain K that contains (t,S,X) (t',S',X')*

- *There is a constant C_3 (possibly dependent of ε) such that:*

$$max(\sigma^2(x), \sigma^2(x'), x^2, x'^2, 1) \leq \frac{C_3}{3}$$

Then the following inequality holds:

$$\begin{aligned} F\left(t', s', x', p', q', A' - D^2 f^\varepsilon(S', X')\right)) - F\left(t, s, x, p, q, A + D^2 f^\varepsilon(s, x)\right) \leq \\ 3C_2^K \alpha \left\|(t - t', s - s', x - x')\right\|^2 + C_2^K \left\|x - x'\right\| q_2 + C_1 \varepsilon + \left|p - p'\right| \\ + C_3 \left\|(x - x_\varepsilon, s - s_\varepsilon)\right\| + C_3 \left\|(x' - x_\varepsilon, s' - s_\varepsilon)\right\| \end{aligned} \tag{4.6}$$

Proof. This proof is an adaptation of example 3.6 in [27]. First, by multiplying inequality (4.5) by $\begin{pmatrix} \Sigma & \Sigma' \end{pmatrix}$ on the left and $\begin{pmatrix} \Sigma \\ \Sigma' \end{pmatrix}$ on the right, one gets:

$$\Sigma A \Sigma - \Sigma' A' \Sigma' \leq 3\alpha \left(\Sigma - \Sigma'\right)^2$$

then, we introduce the symmetric matrices:

$$\begin{aligned} B &= \begin{pmatrix} p + \mu(x) q_2 & 0 \\ 0 & 0 \end{pmatrix} + \Sigma D^2 f^\varepsilon(s, x) \Sigma \\ &= \begin{pmatrix} p + \mu(x) q_2 + \sigma^2(x)(\varepsilon^2 + 3(s - s_\varepsilon)^2) & 0 \\ 0 & max(1, x^2)(3(x - x_\varepsilon)^2) \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} B' &= \begin{pmatrix} p' + \mu(x') q_2' & 0 \\ 0 & 0 \end{pmatrix} - \Sigma D^2 f^\varepsilon(s', x') \Sigma \\ &= \begin{pmatrix} p' + \mu(x') q_2' - \sigma^2(x)(\varepsilon^2 + 3(s' - s_\varepsilon)^2) & 0 \\ 0 & -max(1, x'^2)(3(x' - x_\varepsilon)^2) \end{pmatrix} \end{aligned}$$

And we add $B - B'$ on both sides:

$$\Sigma A \Sigma + B - \Sigma' A' \Sigma' - B' \leq 3\alpha \left(\Sigma - \Sigma'\right)^2 + B - B'$$

and

$$\Sigma A \Sigma + B \leq \Sigma' A' \Sigma' + B' + 3\alpha \left(\Sigma - \Sigma'\right)^2 + B - B'$$

Then, we use the fact that for two symmetric matrices X and Y one has:

$$\lambda^+(X + Y) \leq \lambda^+(X) + \lambda^+(Y)$$

Where λ^+ is the largest eigenvalue of a symmetric matrix. This gives:

$$\begin{aligned} \lambda^+(\Sigma A \Sigma + B) &\leq \lambda^+(\Sigma' A' \Sigma' + B' + 3\alpha \left(\Sigma - \Sigma'\right)^2 + B - B') \\ &\leq \lambda^+(\Sigma' A' \Sigma' + B') + \lambda^+(3\alpha \left(\Sigma - \Sigma'\right)^2) + \lambda^+(B - B') \end{aligned}$$

Thus, knowing that $\lambda^+(X) = -\lambda^-(-X)$ where λ^- is the smallest eigenvalue, we obtain:

$$\lambda^-(-\Sigma' A' \Sigma' - B') - \lambda^-(-\Sigma A \Sigma - B) \leq 3\alpha \lambda^+(\left(\Sigma - \Sigma'\right)^2) + \lambda^+(B - B')$$

By definition (2.10) of the operator F, this is equivalent to:

$$F\left(t', s', x', p, q, A' - D^2 f^\varepsilon(s', x')\right)) - F\left(t, s, x, p, q, A + D^2 f^\varepsilon(s, x)\right) \leq 3\alpha\lambda^+(\left(\Sigma - \Sigma'\right)^2) + \lambda^+(B - B') \tag{4.7}$$

Now, we focus on the right hand side of (4.7) to obtain the result. First we have, as B and B' are diagonal matrices:

$$\begin{aligned}
\lambda^+(B - B') &= \max\left[B_{11} - B'_{11}, B_{22} - B'_{22}\right] \\
&\leq \varepsilon^2(\sigma^2(x) + \sigma^2(x')) + \left|\mu(x) - \mu(x')\right| |q_2| \\
&+ C_3(\|(x - x_\varepsilon, s - s_\varepsilon)\| + \left\|(x' - x_\varepsilon, s' - s_\varepsilon)\right\|) \\
&\leq \varepsilon C_1 + C_2^K \left|x - x'\right| |q_2| + \left|p - p'\right| \\
&+ C_3(\|(x - x_\varepsilon, s - s_\varepsilon)\| + \left\|(x' - x_\varepsilon, s' - s_\varepsilon)\right\|)
\end{aligned}$$

Finally, since Σ is Lipschitz continuous one gets:

$$3\alpha\lambda^+(\left(\Sigma - \Sigma'\right)^2) \leq 3\alpha C_2^K \left\|(t - t', s - s', x - x')\right\|^2$$

Plugging these two inequalities into 4.7, one obtains inequality 4.6, thus proving the lemma. □

4.5 Proof of the comparison principle

Proposition 4.8. *Suppose u is an upper semicontinuous viscosity subsolution of 2.10, bounded from above and w a lower semicontinuous η-strict viscosity supersolution of (2.10) bounded from below. If, furthermore, for some bounded functions g and h:*

$$u(T, ., .) \leq g(.) \leq w(T, ., .) \text{ and } u(., ., 0) \leq h(., .) \leq w(., ., 0)$$

Then $u^*(t, s, x) \leq w_*(t, s, x)$ for all $(t, s, x) \in [O, T] \times [0, +\infty)^2$

Proof. This proof is inspired by [23] and [48]. For $\varepsilon, \alpha > 0$, let $\Phi^{\varepsilon,\alpha}$ be the upper semicontinuous function:

$$\Phi^{\varepsilon,\alpha}(t, t', s, s', x, x') = u(t, s, x) - w(t', s', x') - f_1^\varepsilon(s, s', x, x') + \varepsilon \ln\left(\frac{t}{T}\right) - \alpha\left(d(t - t', s - s', x - x')\right)$$

where

$$d(a, b, c) = \frac{1}{2}\left(a^2 + b^2 + c^2\right)$$

and

$$f_1^\varepsilon(s, s', x, x') = (s + s' + x + x')\varepsilon - (\ln(s) + \ln(s'))\varepsilon^2$$

to simplify, denote

$$f_1^\varepsilon(s, x) = f_1^\varepsilon(s, s, x, x)$$

Next, set:

$$\Phi^\varepsilon(t, S, X) = \Phi^{\varepsilon,\alpha}(t, t, s, s, x, x)$$

As $(u-w)$ is bounded from above by a constant C, one can see from the form of $f(s,s,x,x)$ that the supremum of Φ^ε is attained in $[\exp(\frac{-\beta}{C}),T]\times[\exp(\frac{-\varepsilon}{C}),\frac{\varepsilon}{C}]\times[0,\frac{\varepsilon}{C}]$ which is a compact set. Therefore, Φ^ε is an upper semicontinious function and attains its supremum in a compact set, this supremum is a maximum. It follows that we can find a point $(t_\varepsilon,s_\varepsilon,x_\varepsilon)$ such that:

$$\max_{[O,T]\times[0,+\infty[^2}\Phi^\varepsilon(t,s,x)=\Phi^\varepsilon(t_\varepsilon,s_\varepsilon,x_\varepsilon)$$

Now, there are three possible cases:

- there exist a sequence $\varepsilon_k>0$ such that $\varepsilon_k\to 0$ and $t_{\varepsilon_k}=T$ for every k.
- there exist a sequence $\varepsilon_k>0$ such that $\varepsilon_k\to 0$ and $x_{\varepsilon_k}=0$ for every k.
- there exist a constant $\varepsilon^->0$ such that $x_\varepsilon>0$ and $t_\varepsilon<T$ for all $0<\varepsilon<\varepsilon^-$

Cases 1 and 2: One can prove easily that there is a contradiction if one of the two first cases apply. Indeed, in the first case, one can see that for all ε_k:

$$\begin{aligned}
u(t,s,x)-w(t,s,x) &= \Phi^{\varepsilon_k}(t,s,x)+f^{\varepsilon_k}(s,s,x,x)+\varepsilon_k\ln\left(\frac{t}{T}\right)\\
&\leq \Phi^{\varepsilon_k}(T,s_k,x_k)+f^{\varepsilon_k}(s,s,x,x)-\varepsilon_k\ln\left(\frac{t}{T}\right)\\
&= u(T,s_k,x_k)-w(T,s_k,x_k)+f^{\varepsilon_k}(s,x)-\varepsilon_k\ln\left(\frac{t}{T}\right)\\
&- f^{\varepsilon_k}(s_k,x_k)\\
&\leq u(T,s_k,x_k)-w(T,s_k,x_k)+f^{\varepsilon_k}(s,x)-\varepsilon_k\ln\left(\frac{t}{T}\right)
\end{aligned}$$

because $f^{\varepsilon_k}(s_k,x_k)\geq 0$. Since $u(T,.,.)\leq g(.)\leq w(T,.,.)$ this implies

$$u(t,s,x)-w(t,s,x)\leq f^{\varepsilon_k}(s,x)-\varepsilon_k\ln\left(\frac{t}{T}\right)$$

For all $(t,s,x)\in[O,T]\times[0,+\infty[^2$ hence the proposition is proved by taking $k\to+\infty$. The same kind of proof applies for the second case.

Case 3 This is the technical part. Consider the function:

$$\begin{aligned}
\hat{\Phi}^{\varepsilon,\alpha}(t,t',s,s',x,x') := \Phi^{\varepsilon,\alpha}(t,t',x,s',x,x')-\frac{1}{2}\left[(t-t_\varepsilon)^2+(t'-t_\varepsilon)^2\right]\\
-\frac{1}{4}\left[(s-s_\varepsilon)^4+(s'-s_\varepsilon)^4+(x-x_\varepsilon)^4+(x'-x_\varepsilon)^4\right]
\end{aligned}$$

In the following, we denote $f_2^\varepsilon(s,x)=\frac{1}{4}\left[(s-s_\varepsilon)^4+(x-x_\varepsilon)^4\right]$ and $\hat{\Phi}^\varepsilon(t,s,x):=\hat{\Phi}^{\varepsilon,\alpha}(t,t,s,s,x,x)$. It is clear that

$$\hat{\Phi}^{\varepsilon,\alpha}(t,s,x):=\Phi^{\varepsilon,\alpha}(t,s,x)-(t-t_\varepsilon)^2-2f_2^\varepsilon(s,x)$$

Then for every $\varepsilon>0$, $(t_\varepsilon,s_\varepsilon,x_\varepsilon)$ is a strict maximizer of $\hat{\Phi}^\varepsilon$. Therefore, by lemma 3.1 in [27], for every $\varepsilon<\varepsilon^-$ there exist a sequence $\alpha_k\to+\infty$ and maximizers $(t_{\varepsilon,\alpha},t'_{\varepsilon,\alpha},s_{\varepsilon,\alpha},s'_{\varepsilon,\alpha},x_{\varepsilon,\alpha},x'_{\varepsilon,\alpha})$ of $\hat{\Phi}^{\varepsilon,\alpha}$ such that:

$$(t_{\varepsilon,\alpha},t'_{\varepsilon,\alpha},s_{\varepsilon,\alpha},s'_{\varepsilon,\alpha},x_{\varepsilon,\alpha},x'_{\varepsilon,\alpha})\to(t_\varepsilon,t_\varepsilon,s_\varepsilon,s_\varepsilon,x_\varepsilon,x_\varepsilon)$$

and

$$\alpha_k \left\|(t_{\varepsilon,\alpha} - t'_{\varepsilon,\alpha}, s_{\varepsilon,\alpha} - s'_{\varepsilon,\alpha}, x_{\varepsilon,\alpha} - x'_{\varepsilon,\alpha})\right\|^2 \to 0$$

And, as $(t_\varepsilon, s_\varepsilon, x_\varepsilon)$ are in the interior of the domain for $\varepsilon < \varepsilon^-$, then the maximizers of $\hat{\Phi}^{\varepsilon,\alpha}$ are also in its interior for α_k sufficiently large. With this result, we can apply theorem 3.2 in [27], to the sequence of local maxima. We obtain that, for sufficiently large α_k, there exists two symmetric matrices $A_k, A'_k \in s^2$ such that:

$$\begin{aligned}
&(A_k, A'_k) \text{ satisfies 4.5},\\
&\left(p_k, q_k + D\left(f_1^\varepsilon + f_2^\varepsilon\right)(s_k, x_k), A_k + D^2\left(f_1^\varepsilon + f_2^\varepsilon\right)(s_k, x_k)\right) \in J^{2,-}u(t_k, s_k, x_k),\\
&\left(p'_k, q_k - D\left(f_1^\varepsilon + f_2^\varepsilon\right)(s'_k, x'_k), A'_k - D^2\left(f_1^\varepsilon + f_2^\varepsilon\right)(s'_k, x'_k)\right) \in J^{2,+}w(t'_k, s'_k, x'_k)
\end{aligned}$$

where, by taking $f = f_1 + f_2$

$$\begin{aligned}
p_k &= \alpha_k\left(t_k - t'_k\right) + (t_k - t_\varepsilon)\\
p'_k &= \alpha_k\left(t_k - t'_k\right) - \left(t'_k - t_\varepsilon\right)\\
q_k &= \alpha_k \begin{pmatrix} s_k - s'_k \\ x_k - x'_k \end{pmatrix}\\
D^2 f^\varepsilon(s_k, x_k) &= \begin{pmatrix} \frac{\varepsilon^2}{s_k^2} + 3(s_k - s_\varepsilon)^2 & 0 \\ 0 & 3(x_k - x_\varepsilon)^2 \end{pmatrix}
\end{aligned}$$

and with $J^{2,+}w(t'_k, s'_k, x'_k)$ and $J^{2,-}u(t_k, s_k, x_k)$ are, as in [27], the closed inferior and superior semijets of w and u respectively. Then by the definition of viscosity subsolutions and strict supersolutions we obtain:

$$F\left(t_k, s_k, x_k, p_k, A_k + D^2 f^\varepsilon(s_k, x_k)\right) \leq 0$$

and

$$F\left(t'_k, s'_k, x'_k, p'_k, A'_k - D^2 f^\varepsilon(s'_k, x'_k)\right) > \eta$$

Combining these two inequalities, one gets:

$$F\left(t'_k, s'_k, x'_k, p'_k, q'_k, A'_k - D^2 f^\varepsilon(s'_k, x'_k)\right) - F\left(t_k, s_k, x_k, p_k, q_k, A_k + D^2 f^\varepsilon(s_k, x_k)\right) > \eta \quad (4.8)$$

On the other hand, since the maximum point $t_\varepsilon, s_\varepsilon, x_\varepsilon$ is attained in $[\exp(\frac{-\beta}{C}), T] \times [\exp(\frac{-\varepsilon}{C}), \frac{\varepsilon}{C}] \times]0, \frac{\varepsilon}{C}]$, and since (t_k, s_k, x_k) and (t'_k, s'_k, x'_k) converge to this point, the local Lipschitz condition (2.2) proves the existence of the constants $C_1, C_2^\varepsilon, C_3$, of lemma 4.2, independent of k provided it is sufficiently large, and with C_1 independent of ε provided it is sufficiently small. Now we can apply it to obtain gives the inequality:

$$\begin{aligned}
F\left(t'_k, s'_k, x'_k, p_k, A'_k - \varepsilon Dl(s'_k, x'_k)\right) - F\left(t_k, s_k, x_k, p_k, A_k + \varepsilon Dl(s_k, x_k)\right) \leq&\\
C_2^\varepsilon \alpha_k \left\|(t_{\varepsilon,\alpha} - t'_{\varepsilon,\alpha}, s_{\varepsilon,\alpha} - s'_{\varepsilon,\alpha}, x_{\varepsilon,\alpha} - x'_{\varepsilon,\alpha})\right\|^2 + C_1\varepsilon&\\
+ C_3(\|(x - x_\varepsilon, s - s_\varepsilon)\| + \left\|(x' - x_\varepsilon, s' - s_\varepsilon)\right\|)& \quad (4.9)
\end{aligned}$$

The right hand side of (4.9) tends to $C_1\varepsilon$ when k tends to infinity, and by sending ε to zero, this contradicts (4.8), thus proving the comparison result. □

5 Dual representation

In this section, we give a dual expectation representation of the super-replication problem. The dual maximization problem is done over all volatilities of X and all possible correlations between X and S. This kind of duality was first introduced in [65] for one-dimensional processes. First we define the value function $\tilde{v}$ of the dual problem:

$$\tilde{v}(t,s,x) = \sup_{(\rho,\xi)\in\mathcal{U}} E\left[g\left(S^{\rho,\xi}_{t,s,x}(T)\right)\right] \tag{5.1}$$

Where $\mathcal{U}$ is the set of all almost-surely bounded progressively measurable processes taking values in $[-1;1]\times[0;+\infty)$:

$$\mathcal{U} = \left\{(\rho,\xi) \text{ valued in } [-1,1]\times[0,+\infty) \text{ and prog. measurable} | \int_0^T \xi_t^2 dt < +\infty\right\}$$

And the process $S^{\rho,\xi}_{t,s}$ is defined for $u \geq t$ by the dynamics:

$$\begin{cases} S^{\rho,\xi}_{t,s,x}(t) = s \quad \text{and} \quad X^{\rho,\xi}_{tx}(t) = x \\ dS^{\rho,\xi}_{t,s,x}(u) = \sigma\left(t, X^{\rho,\xi}_x(u)\right) S^{\rho,\xi}_{t,s,x} dW^1(u) \\ dX^{\rho,\xi}_x(u) = -\mu\left(t, X^{\rho,\xi}_x(u)\right) du + \xi_u X^{\rho,\xi}_x(u)\, dW^2(u) \\ \left\langle dW^1(u), dW^2(u)\right\rangle = \rho_u \end{cases}$$

The main goal of this section is to prove that $\tilde{v}$ is also solution of the primal super-replication problem. In other words that $\tilde{v} = v$. First, we have to prove that the two functions verify the same equation. Then, we prove that the two function have the same boundary conditions. We then conclude by the comparison theorem.

Proposition 5.9. *$\tilde{v}$ is a viscosity supersolution of equation (2.10) on* $(0;T)\times(0,+\infty)^2$.

Proof. This is a classical proof in the optimal control theory, see [59], chapter 4, for details. That framework applies to one-dimensional problems, but there is no difficulty in extending them to the multidimensional case. Hence the viscosity sub and supersolution characterization in terms of Hamiltonian will be admitted, and we will focus on the equivalence between the classical characterization and equation (2.10). The Hamiltonian of the problem is:

$$\begin{aligned} H\left(s,x,v,Dv,D^2v\right) = \inf_{\xi,\rho} \Bigg\{ & \frac{\partial v}{\partial x}\mu(t,X) - \frac{1}{2}\xi^2\frac{\partial^2 v}{\partial x^2} - \frac{1}{2}\sigma^2(t,x)s^2\frac{\partial^2 v}{\partial s^2} \\ & -\rho\xi\sigma(t,x)s\frac{\partial v}{\partial s\partial x}\Bigg\} \end{aligned}$$

And by classical techniques one can show that:

$$-\frac{\partial v}{\partial t} + H\left(s,x,v,Dv,D^2v\right) \geq 0 \tag{5.2}$$

in the viscosity sense. Indeed, the Hamiltonian H is smooth, unless it takes infinite negative values. In order to prove that the continuous operator $F\left(S,X,\tilde{v},D\tilde{v},D^2\tilde{v}\right)$ of

(2.10) is such that F and $H - \frac{\partial}{\partial t}$ always have the same sign, the next step is to explicitly solve the Hamiltonian H. Writing, for any vector b and any 2×2 symmetric matrix A:

$$H(S, X, v, b, A) = \inf_{\xi,\rho} \left\{ b_2\mu(t, X) - \frac{1}{2}\xi^2 A_{22} - \frac{1}{2}\sigma^2(t, X)S^2 A_{11} - \rho\xi\sigma(t, X)SA_{12} \right\}$$

By elementary techniques, the minimization of H over ρ and ξ gives:

$$\left\{ \begin{array}{c} H(S, X, v, b, A) = -\infty \text{ if } A_{22} > 0 \\ H(S, X, v, b, A) = -\infty \text{ if } A_{22} = 0 \text{ and } A_{12} \neq 0 \\ H(S, X, v, b, A) = b_2\mu(t, X) - \frac{1}{2}\sigma^2(t, X)S^2 A_{11} \text{ if } A_{22} = 0 \text{ and } A_{12} = 0 \\ H(S, X, v, b, A) = \frac{1}{2}\sigma^2(t, X)S^2 \left(\frac{A_{12}^2}{A_{22}} - A_{11} \right) + b_2\mu(t, X) \text{ otherwise} \end{array} \right. \tag{5.3}$$

In addition, the operator F is positive if and only if the matrix $J\left(t, s, x, Du, D^2u\right)$ defined in (2.10) is positive, that is, if and only if the two diagonal terms J_{11} and J_{22} and the determinant of J are positive. Clearly, by (5.3), F is positive if and only if H is positive. Hence $\tilde{v}$ is a viscosity supersolution of (2.10). □

Now, we concentrate on the subsolution property:

Proposition 5.10. *$\tilde{v}$ is a viscosity subsolution of equation (2.10) on $(0;T) \times (0, +\infty)^2$.*

Proof. Let $\varphi \in C^2((0;T) \times (0, +\infty)^2)$ a test function such that

$$0 = (v^* - \varphi)\left(\bar{t}, \overline{S}, \overline{X}\right) = \max_{(t,S,X)\in(0;T)\times(0,+\infty)^2} (v^* - \varphi)(t, S, X)$$

for some $(\bar{t}, \overline{S}, \overline{X}) \in (0;T) \times (0, \infty)^2$. Suppose that, on the contrary,

$$F\left(t, S, X, \varphi, D\varphi, D^2\varphi\right) > 0$$

Hence

$$-\frac{\partial\tilde{v}}{\partial t} + H\left(t, S, X, \varphi, D\varphi, D^2\varphi\right) > 0$$

considering (5.3), as H is continuous in the interior of the domain delimited by $F > 0$, one can find a contradiction with a classical dynamic programming argument which can be found in [59] for instance. □

In order to apply the uniqueness proposition 4.7, it would remains to verify that the value function $\tilde{v}$ of the dual problem has the same boundary conditions as v. There are two parts in this question: The study for $t \to T$ and for $X \to 0$. Ideally, we would prove the two following propositions directly. However this may be quite difficult, and they will be demonstrated indirectly along the lines of the proof of proposition 5.13. Let us begin by the first limit:

Proposition 5.11. *The value function $\tilde{v}$ of problem (5.1) extends continuously to a function $\widehat{\tilde{v}}$ on $(0;T] \times (0, +\infty)^2$ satisfying the terminal condition:*

$$\widehat{\tilde{v}}(T, S, X) = g(S)$$

Moreover, we need the condition near $X = 0$

Proposition 5.12. *The function* $\widehat{\tilde{v}}$ *extends continuously to a function* $\overline{\tilde{v}}$ *on* $(0;T] \times (\mathbb{R}_+)^2$ *satisfying the boundary condition:*

$$\overline{\tilde{v}}(t,S,0) = g(S)$$

The study of the behavior of $\tilde{v}$ near $X = 0$ requires several steps. We define the auxiliary value function:

$$\tilde{v}_C = \sup_{(\xi,\rho)\in U_C} E\left[g\left(S^{x,\xi,\rho}_{t,T}\right)\right]$$

Where $U_C = \{(\xi,\rho) \in U | \xi \le C\}$. The preliminary goal is to prove the following technical lemma :

Lemma 5.3. *For any* $C \ge 0$ *there exists two constants* C_1 *and* C_2 *independent of* $(x,t,u) \in \mathbb{R} \times [0;T]^2$ *with* $t < u$ *such that for any adapted processes* $(\xi,\rho) \in [0;C] \times [-1,1]$ *one has:*

$$E\left[\left(X^{x,\xi,\rho}_{t,u}\right)^2\right] \le x^2 + C_1 x^2 \int_t^u e^{C_1(u-s)} ds \tag{5.4}$$

$$E\left[X^{x,\xi,\rho}_{t,u}\right] \le x + C_2 x \int_t^u e^{C_2(u-s)} ds \tag{5.5}$$

Proof. We use a similar procedure as in [59]. By Itô's formula, one has, for any stopping time τ, and any $0 \le t \le s, x, (\xi,\rho) \in U_C$:

$$\left(X^{x,\xi,\rho}_{t,u\wedge\tau}\right)^2 = x^2 + \int_t^{u\wedge\tau} \left[-2X^{x,\xi,\rho}_{t,s}\mu\left(s, X^{x,\xi,\rho}_{t,s}\right) + \left(\xi X^{x,\xi,\rho}_{t,s}\right)^2\right] ds + \int_t^{u\wedge\tau} \xi \left(X^{x,\xi,\rho}_{t,s}\right)^2 dW^2(s)$$

choosing a sequence of stopping times: $\tau_n = \inf\left\{s \ge t : \xi\left(X^{x,\xi,\rho}_{t,s}\right)^2 \ge n\right\}$, which tends a.s. to infinity when $n \to +\infty$, when have for fixed n :

$$E\left[\int_t^{u\wedge\tau_n} \xi\left(X^{x,\xi,\rho}_{t,s}\right)^2 dW^2(s)\right] = 0$$

Then, using the linear growth coefficient K of μ and the bound C of ξ:

$$\begin{aligned}
E\left[\left(X^{x,\xi,\rho}_{t,u\wedge\tau_n}\right)^2\right] &= x^2 + E\left(\int_t^{u\wedge\tau_n}\left[-2X^{x,\xi,\rho}_{t,s}\mu\left(s, X^{x,\xi,\rho}_{t,s}\right) + \left(\xi X^{x,\xi,\rho}_{t,s}\right)^2\right] ds\right) \\
&\le x^2 + E\left(\int_t^{u\wedge\tau_n}\left[2K\left(X^{x,\xi,\rho}_{t,s}\right)^2 + C\left(X^{x,\xi,\rho}_{t,s}\right)^2\right] ds\right) \\
&\le x^2 + (2K+C)\, E\left(\int_t^{u\wedge\tau_n}\left[\left(X^{x,\xi,\rho}_{t,s}\right)^2\right] ds\right)
\end{aligned}$$

By Gronwall's lemma, writing $C_1 = 2K + C$:

$$E\left[\left(X^{x,\xi,\rho}_{t,u\wedge\tau_n}\right)^2\right] \le x^2 + C_1 x^2 \int_t^{u\wedge\tau_n} e^{C_1(u\wedge\tau_n - s)} ds$$

Finally, using Fatou's lemma and letting $n \to +\infty$

$$E\left[\left(X_{t,u}^{x,\xi,\rho}\right)^2\right] \leq x^2 + C_1 x^2 \int_t^u e^{C_1(u-s)} ds$$

Moreover, with the same arguments, one can prove that there exists a constant C_2 independent of x, ξ, ρ, t, u such that:

$$E\left[\left(X_{t,u}^{x,\xi,\rho}\right)\right] \leq x + C_2 x \int_t^u e^{C_2(u-s)} ds$$

□

Now we have the tools to prove the convergence results for v_C.

Lemma 5.4. *The terminal condition of $\tilde{v}_C$ is:*

$$\lim_{(t,s',x')\to(T,s,x)} \tilde{v}_C(t,s,x) = g(s) \text{ for any } (s,x,C) \in \mathbb{R}_+^3 \tag{5.6}$$

Furthermore, the lateral condition of $\tilde{v}_C$ is:

$$\lim_{(t',s',x)\to(t,s,0)} \tilde{v}_C(t,s,x) = g(s) \text{ for any } (t,s,C) \in [0,T] \times \mathbb{R}_+^2 \tag{5.7}$$

Proof. For sake of conciseness we prove the two propositions at the same time. Since function g is bounded, function $\widetilde{v}_C$ has the same bounds. Hence, if for a given point (t,s,x) and for any sequence $(t_n, s_n, x_n) \to (t,s,x)$, $\widetilde{v}_C(t_n, s_n, x_n)$ admits $g(s)$ as an accumulation point, then $\widetilde{v}_C$ is continuous at (t,s,x) and equal to $g(s)$. Therefore, we have to prove this claim at points of type (T,s,x) and $(t,s,0)$. Choose a sequence $(t_n, s_n, x_n) \in [0,T] \times \mathbb{R}_+^2$ converging to a given (T,s,x) or $(t,s,0)$. Then, by definition of the value function $\widetilde{v}_C$ there exists a sequence of controls $(\xi_n, \rho_n) \in U_C$ such that, denoting S^n as S^{ξ_n,ρ_n}:

$$E\left(g(S^n_{t_n,s_n,x_n})\right) \leq \widetilde{v}_C(t_n,s_n,x_n) \leq E\left(g(S^n_{t_n,s_n,x_n})\right) + \frac{1}{n} \tag{5.8}$$

Now, we get to the convergence of both sides. We use the Doleans exponential formula:

$$S^n_{t_n,s_n,x_n} = s_n \exp\left(\int_{t_n}^T \sigma\left(u, X^n_{t_n,x_n}(u)\right) dW_u + \frac{1}{2}\int_{t_n}^T \sigma^2\left(u, X^n_{t_n,x_n}(u)\right) du\right)$$

One gets, taking the logarithm (will not work with $s = 0$ but then the proof is trivial):

$$\begin{aligned}\ln(S^n_{t_n,s_n,x_n}) - \ln(s) =& \ln(s_n) - \ln(s) + \int_{t_n}^T \sigma\left(u, X^n_{t_n,x_n}(u)\right) dW_u \\ &+ \frac{1}{2}\int_{t_n}^T \sigma^2\left(u, X^n_{t_n,x_n}(u)\right) du\end{aligned}$$

Taking the square of this equality one gets:

$$\begin{aligned}\left(\ln(S^n_{t_n,s_n,x_n}) - \ln(s)\right)^2 = (\ln(s_n) - \ln(s))^2 &+ \left(\int_{t_n}^T \sigma\left(u, X^n_{t_n,x_n}(u)\right) dW_u\right)^2 \\ &+ \left(\int_{t_n}^T \frac{1}{2}\int_{t_n}^T \sigma^2\left(u, X^n_{t_n,x_n}(u)\right) du\right)^2\end{aligned}$$

Using (2.6) together with (5.4) gives:

$$E\left[\int_{t_n}^{T} \sigma^2\left(u, X^n_{t_n,x_n}(u)\right) du\right] \leq (T-t_n)Cx$$

$$E\left[\int_{t_n}^{T} \sigma^2\left(u, X^n_{t_n,x_n}(u)\right) du\right]^2 \leq (T-t_n)Cx^2$$

And where C is independent of n. Hence, in the context of lemma 5.4 we get:

$$\lim_{n\to+\infty} \ln(S^n_{t_n,s_n,x_n}) = \ln(s) \text{ in } L^2$$

Then, there exists a subsequence n_k satisfying:

$$\lim_{k\to+\infty} \ln(S^{n_k}_{t_{n_k},s_{n_k},x_{n_k}}) = \ln(s) \text{ almost surely}$$

And as g is continuous and bounded, we conclude, by the dominated convergence theorem:

$$\lim_{k\to+\infty} E\left(g(S^{n_k}_{t_{n_k},s_{n_k},x_{n_k}})\right) = E\left(g(\lim_{k\to+\infty} S^{n_k}_{t_{n_k},s_{n_k},x_{n_k}})\right) = g(s)$$

Remembering inequalities (5.8), we get that $g(s)$ is an accumulation point of $\widetilde{v}_C(t_n, s_n, x_n)$ then the proof is complete. □

Now we are in position to prove the main result of this section:

Proposition 5.13. *The value function of the primal problem an the dual problem are the same. In other words:*

$$\tilde{v} = v$$

Proof. We use the fact that $\widetilde{v}^*_C$ is a viscosity subsolution of the following equation (5.9), studied in [12]. We will give no proof of this claim, see that paper for detail

$$\inf_{\rho\xi\in[-C;C]} \left\{-\mathcal{G}^{\rho\xi}\varphi\right\} \geq 0 \text{ where} \tag{5.9}$$

$$\mathcal{G}^{\rho\xi}\varphi = \frac{\partial\varphi}{\partial t} + \frac{1}{2}\begin{pmatrix} S\sigma(t,X) & \rho\xi \end{pmatrix} \begin{pmatrix} \frac{\partial^2\varphi}{\partial S^2} & \frac{\partial^2\varphi}{\partial S\partial X} \\ \frac{\partial^2\varphi}{\partial S\partial X} & \frac{\partial^2\varphi}{\partial X^2} \end{pmatrix} \begin{pmatrix} S\sigma(t,X) \\ \rho\xi \end{pmatrix} \tag{5.10}$$

And, because the negativity of matrix H^1 follows trivially from the negativity of $\mathcal{G}^{\rho\xi}\varphi$ (which is an minimum of the quadratic form defined by H over vectors of form $\begin{pmatrix} 1 \\ \rho\xi \end{pmatrix}$), we obtain that $\tilde{v}^*_C$ is a viscosity subsolution of equation (2.10) in the viscosity sense. Together with terminal conditions (5.6) and (5.7), we obtain:

$$\tilde{v}_C \leq v$$

Letting $C \to +\infty$ we get

$$\tilde{v} \leq v$$

Furthermore as

$$\tilde{v}_0 \leq \tilde{v}$$

and $\tilde{v}_0$ is a Black-Scholes price with deterministic volatility, we have, for any $(t,s) \in [0,T] \times \mathbb{R}$

$$\lim_{x\to 0} \tilde{v}_0\left(t,s,x\right) = g\left(s\right)$$

Knowing the boundary condition of v by propositions 3.5 and 3.6, and as $\tilde{v}_0 \leq \tilde{v} \leq v$ we have :

$$\lim_{x\to 0} \tilde{v}\left(t,s,x\right) = \lim_{x\to 0} v\left(t,s,x\right) = g\left(s\right)$$

Hence, we can conclude by the comparison principle of proposition 4.7 that:

$$\tilde{v} = v$$

□

6 Law of the iterated logarithm for some double stochastic integrals

Here, we prove the lemma which we use in the demonstration of the super-solution property. In particular, we have to show that if the matrix $\Gamma - D^2\phi$ was constant, it would have to be symmetric and positive in order to satisfy relation (3.10). Then, corollary 3.8 in [24], proves, under some regularity assumptions, that if $\Gamma - D^2\phi$ changes over time, then a necessary and sufficient condition for (3.10) to hold is that $\left(\Gamma - D^2\phi\right)(0)$ is symmetric positive. As the symmetric case is studied in [42] we will deal with the non-symmetric case. The key issue is to estimate the limit of processes written as:

$$\liminf_{t\to 0} \frac{\int_0^t \left(\int_0^u \begin{bmatrix} a & 1 \\ -1 & a \end{bmatrix} dW_v\right)^T dW_u}{t \log\log \frac{1}{t}}$$

Integrating the diagonal part, one obtains:

$$\frac{a}{2}\left(W_{1t}^2 + W_{2t}^2\right) - at + \int_0^t W_{1u}dW_{2u} - W_{2u}dW_{1u}$$

Hence, it is sufficient to study the limit of:

$$Z_t = \frac{a}{2}\left(W_{1t}^2 + W_{2t}^2\right) + \int_0^t W_{1u}dW_{2u} - W_{2u}dW_{1u}$$

The process L_t is called the Levy area:

$$L_t = \int_0^t W_{1u}dW_{2u} - W_{2u}dW_{1u}$$

First, let us see how this study will enable us to solve the consider problem. That is, we must find a result like the main theorem of [24] that embeds the case of nonsymmetric matrices:

Proposition 6.14. *Let $M(t)$ be an $\mathbb{R}^2$-valued martingale defined for any $t > 0$ by:*

$$M(t) = \int_0^t m(r)dW_r$$

where $m(t)$ is a $\mathcal{M}^2$ valued, $\mathbb{F}$ progressively measurable process such that, for any $t > 0$:

$$\int_0^t |m(r)|^2 \, dr < +\infty$$

Let $b(t)$ be a bounded, $\mathcal{M}^2$ valued, $\mathbb{F}$ progressively measurable process,and assume there exist a random variable $\varepsilon > 0$ such that almost surely:

$$\int_0^t |m(r) - m(0)|^2 \, dr = O(t^{1+\varepsilon}) \text{ and } \int_0^t |b(r) - b(0)|^2 \, dr = O(t^{1+\varepsilon}) \tag{6.1}$$

For $t \to 0^+$. Then:

$$\liminf_{t\to 0^+} \frac{1}{t \log\log \frac{1}{t}} \int_0^t \left(\int_0^r b(u)dM_u \right)^T dM_r \geq 0 \text{ if and only if } b(0) \text{ is symmetric positive}$$

Proof. The proof is an extension of proof of theorem 3.3 in [24]. If $b(0)$ is symmetric, the proof is already done in corollary 3.7 of that paper. So, suppose that $b(0)$ is not symmetric. One can decompose the integral into:

$$\int_0^t \left(\int_0^r b(u)dM_u \right)^T dM_r = \int_0^t \left(\int_0^r m(0)^T b(0)m(0)dW_u \right)^T dW_r + R_1(t) + R_2(t)$$

where:

$$R_1(t) = \int_0^t \left(\int_0^r b(u)[m(u) - m(0)]dW_u \right)^T m(0)dW_r$$
$$R_2(t) = \int_0^t \left(\int_0^r b(u)m(u)dW_u \right)^T [m(r) - m(0)]dW_r$$
$$R_3(t) = \int_0^t \left(\int_0^r [b(u) - b(0)]m(0)dW_u \right)^T m(0)dW_r$$

in [24] it is shown that assumption (6.1) gives:

$$\lim_{t\to 0} \frac{R_1(t)}{t \log\log \frac{1}{t}} = \lim_{t\to 0} \frac{R_2(t)}{t \log\log \frac{1}{t}} = \lim_{t\to 0} \frac{R_3(t)}{t \log\log \frac{1}{t}} = 0$$

Hence, one only has to study the behavior of $\int_0^t \left(\int_0^r m(0)^T b(0)m(0)dW_u \right)^T dW_r$. We denote $m(0)^T b(0)m(0) = c$. Next we decompose c into a symmetric part c_1 and a skew-symmetric part c_2. We use a base $\tilde{W}$ for W where c_1 is diagonal. In this base:

$$c_1 = \begin{pmatrix} \lambda^+ & 0 \\ 0 & \lambda^- \end{pmatrix} \text{ and } c_2 = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$$

Where $\lambda^+ > \lambda^-$. We define $c_3 = \begin{pmatrix} 0 & 0 \\ 0 & \lambda^- - \lambda^+ \end{pmatrix}$. Then we get:

$$\begin{aligned}\int_0^t \left(\int_0^r c dW_u\right)^T dW_r &= \int_0^t \left(\int_0^r (c_1+c_2) dW_u\right)^T dW_r \\ &= \int_0^t \left(\int_0^r (c_2+\lambda^+ I_2 + c_3) dW_u\right)^T dW_r\end{aligned}$$

Now, as $a \neq 0$ using lemma 6.10 we get that:

$$\liminf_{t\to 0} \frac{1}{t\log\log\frac{1}{t}} \int_0^t \left(\int_0^r (c_2+\lambda^+ I_2) dW_u\right)^T dW_r =$$
$$-|a| \limsup_{t\to 0} \frac{1}{t\log\log\frac{1}{t}} \int_0^t \left(\int_0^r (-\frac{1}{|a|}c_2 - \frac{\lambda^+}{|a|} I_2) dW_u\right)^T dW_r < 0$$

And as

$$\int_0^t \left(\int_0^r (c_3) dW_u\right)^T dW_r = (\lambda^- - \lambda^+)\frac{(\widetilde{W_t^2})^2 - t}{2}$$

and

$$\liminf_{t\to 0} \frac{1}{t\log\log\frac{1}{t}} (\lambda^- - \lambda^+)\frac{(\widetilde{W_t^2})^2 - t}{2} < 0$$

Then the proof is over. □

6.1 Density of the considered process

Now let us study the probability density of the process Z. It is given in this lemma:

Lemma 6.5. *For every $t > 0$, the probability density of the random variable Z_t is:*

$$\varphi(z_t) = \frac{\exp\left(\frac{z}{t}\arctan(a)\right)}{2\rho t\left[ch\left(\frac{\pi z}{2t}\right)\right]}$$

Proof. To begin, we use the Levy formula (see [50]):

$$E\left[e^{i\lambda L_t} \middle| \sqrt{W_{1t}^2 + W_{2t}^2} = x\right] = \frac{t\lambda}{sh(t\lambda)} \exp\left[\frac{|x|^2}{2t}(1 - t\lambda\coth(t\lambda))\right]$$

As the expectation is conditional to $x^2 = W_{1t}^2 + W_{2t}^2$, multiplying it by a function of the conditioner, one obtains:

$$\begin{aligned}E\left[e^{i\lambda\left(L_t + \frac{a}{2}x^2\right)} \middle| \sqrt{W_{1t}^2 + W_{2t}^2} = x\right] &= \frac{t\lambda}{sh(t\lambda)} \exp\left[\frac{|x|^2}{2t}(1 - t\lambda\coth(t\lambda)) + i\lambda\frac{a}{2}x^2\right] \\ &= \frac{t\lambda}{sh(t\lambda)} \exp\left[\frac{|x|^2}{2t}(1 - t\lambda\coth(t\lambda) + t\lambda a i)\right]\end{aligned}$$

As x^2 is the sum of two squared gaussian variables, it is distributed accordingly a two dimensional chi-squared law. Therefore, its probability density is:

$$d\mathbb{P}\left(x^2\right) = \frac{\exp\left(-\frac{x^2}{2t}\right)}{2t} d\left(x^2\right)$$

Thus, integrating expression (6.2) with respect to x^2:

$$\begin{aligned}
E\left[e^{i\lambda\left(L_t+\frac{a}{2}x^2\right)}\right] &= \int_0^{+\infty} E\left[e^{i\lambda\left(L_t+\frac{a}{2}x^2\right)} \middle| \sqrt{W_{1t}^2+W_{2t}^2}=x\right] d\mathbb{P}\left(x^2\right) \\
&= \int_0^{+\infty} \frac{\lambda}{2sh\left(t\lambda\right)} \exp\left[\frac{x^2}{2t}\left(-t\lambda\coth\left(t\lambda\right)+t\lambda ai\right)\right] d\left(x^2\right) \\
&= \left[\frac{1}{sh\left(t\lambda\right)\left(-\coth\left(t\lambda\right)+ai\right)} \exp\left[\frac{x^2}{2t}\left(-t\lambda\coth\left(t\lambda\right)+t\lambda ai\right)\right]\right]_0^{+\infty} \\
&= \frac{1}{ch\left(t\lambda\right)-ai\left[sh\left(t\lambda\right)\right]}
\end{aligned}$$

Then, manipulating the expression to obtain a canonical form, one derives, defining $\rho = \sqrt{1+a^2}$ and $\theta = \arg\left(1-ai\right) = \arctan\left(-a\right)$:

$$\begin{aligned}
E\left[e^{i\lambda\left(L_t+\frac{a}{2}x^2\right)}\right] &= \frac{2}{e^{t\lambda}+e^{-t\lambda}-aie^{t\lambda}+aie^{-t\lambda}} \\
&= \frac{2}{\left(1-ai\right)e^{t\lambda}+\left(1+ai\right)e^{-t\lambda}} \\
&= \frac{2}{\rho\left[e^{t\lambda+i\theta}+e^{-t\lambda-i\theta}\right]} \\
&= \left(\frac{2}{\rho e^{i\theta}}\right)\frac{e^{t\lambda}}{e^{-2i\theta}+e^{2t\lambda}}
\end{aligned}$$

To derive the probability density φ of the random variable $z_t = L_t + \frac{a}{2}x^2$, we must calculate the inverse Fourier transform of this function. Defining $\lambda' = -2t\lambda$ and $y = \frac{z}{2t}$ a change of variables gives:

$$\begin{aligned}
\varphi\left(z\right) &= \left(\frac{2}{2\pi\rho e^{i\theta}}\right)\int_{-\infty}^{+\infty} \frac{e^{t\lambda}e^{-i\lambda z}}{e^{-2i\theta}+e^{2t\lambda}} d\lambda \\
\varphi\left(y\right) &= \frac{2e^{-i\theta}}{4\rho t\pi}\int_{-\infty}^{+\infty} \frac{e^{-\frac{\lambda'}{2}}e^{iy\lambda'}}{e^{-2i\theta}+e^{-\lambda'}} d\lambda'
\end{aligned}$$

This Fourier transform can be found in [37]. Its inverse is:

$$\begin{aligned}
\varphi\left(y\right) &= \frac{2e^{i\theta}}{4\rho t}\frac{e^{-i\theta-2y\theta}}{\sin\left(\frac{\pi}{2}-i\pi y\right)} \\
&= \frac{e^{-2y\theta}}{2\rho t\sin\left(\frac{\pi}{2}-i\pi y\right)} \\
&= \frac{e^{-2y\theta}}{2\rho t\left[ch\left(\pi y\right)\right]}
\end{aligned}$$

Hence we found the density of $z_t = L_t + \frac{a}{2}x^2$

$$\varphi\left(z_t\right) = \frac{\exp\left(-\frac{z\theta}{t}\right)}{2\rho t\left[ch\left(\frac{\pi z}{2t}\right)\right]}$$

$$\varphi\left(z_t\right) = \frac{\exp\left(\frac{z}{t}\arctan(a)\right)}{2\rho t\left[ch\left(\frac{\pi z}{2t}\right)\right]}$$

$\square$

This density function enables us to derive an upper bound for the cumulative distribution of Z_t:

Lemma 6.6. *For given $t > 0$ and $z > 0$, the probability $\mathbb{P}(Z_t > z)$ is majored by:*

$$\mathbb{P}(Z_t > z) \geq 1 - \frac{1}{2\rho}\frac{\exp\left(\left(\arctan(a) - \frac{\pi}{2}\right)\frac{Z}{t}\right)}{\frac{\pi}{2} - \arctan(a)}$$

Proof. Integration of φ gives the probability of Z to be above a given z:

$$\begin{aligned}
F\left(Z\right) &= \int_Z^{+\infty}\frac{\exp\left(\frac{z}{t}\arctan(a)\right)}{2\rho t\left[ch\left(\frac{\pi z}{2t}\right)\right]}dz \\
&= \frac{1}{\rho t}\int_Z^{+\infty}\frac{\exp\left(\frac{z}{t}\arctan(a)\right)}{e^{\frac{\pi z}{2t}} + e^{-\frac{\pi z}{2t}}}dz \\
&\geq \frac{1}{2\rho t}\int_Z^{+\infty}\frac{\exp\left(\frac{z}{t}\arctan(a)\right)}{e^{\frac{\pi z}{2t}}}dz \\
&\geq \frac{1}{2\rho t}\int_Z^{+\infty}\exp\left(\frac{z}{t}\arctan(a) - \frac{\pi z}{2t}\right)dz \\
&\geq \frac{1}{2\rho}\frac{\exp\left(\left(\arctan(a) - \frac{\pi}{2}\right)\frac{Z}{t}\right)}{\frac{\pi}{2} - \arctan(a)}
\end{aligned}$$

Where we use the fact that $z > 0$, hence that $e^{\frac{\pi z}{2t}} + e^{\frac{-\pi z}{2t}} \leq 2e^{\frac{\pi z}{2t}}$. This is the only approximation in this formula, and one can see that the error is less than a factor 2. $\square$

6.2 Approximating the Laplace transform

The Laplace transform of the considered process $Z_t = L_t + \frac{a}{2}\left(W_{1t}^2 + W_{2t}^2\right)$ is defined as follows:

$$\Psi\left(c\right) = E\left(\exp\left(cz_t\right)\right)$$

As one knows the density of Z_t, the following formula is straightforward:

$$\Psi\left(c\right) = \int_{-\infty}^{+\infty}\frac{\exp\left(\left(c - \frac{\theta}{t}\right)z\right)}{2\rho t\left[ch\left(\frac{\pi z}{2t}\right)\right]}dz$$

On can see that $\Psi(c) < +\infty$ iff $-\frac{\pi}{2t} + \frac{\theta}{t} < c < \frac{\pi}{2t} + \frac{\theta}{t}$. If this holds true, then one can dominate the Laplace transform with:

$$\begin{aligned}\Psi(c) &\leq \frac{1}{\rho t}\left[\int_0^{+\infty} \exp\left(\left(c - \frac{\theta}{t} - \frac{\pi}{2t}\right) z\right) dz + \int_{-\infty}^0 \exp\left(\left(c - \frac{\theta}{t} + \frac{\pi}{2t}\right) z\right) dz\right] \\ &\leq \frac{1}{\rho t}\left[\frac{1}{\frac{\theta}{t} + \frac{\pi}{2t} - c} + \frac{1}{c - \frac{\theta}{t} + \frac{\pi}{2t}}\right] \end{aligned} \tag{6.2}$$

6.3 Proof of the law of the iterated logarithm

We study the superior limit of the process

$$Z_t = \frac{a}{2}\left(W_{1t}^2 + W_{2t}^2\right) + \int_0^t W_{1u} dW_{2u} - W_{2u} dW_{1u}$$

In the almost sure sense, when t goes to 0. The Laplace transform of Z_t can be used to show the first estimate:

Lemma 6.7. *The process Z is such that*

$$\lim_{t \to 0^+} \sup \frac{Z_t}{t \log\log \frac{1}{t}} \leq \frac{1}{\frac{\pi}{2} - \arctan(a)}$$

In the almost sure sense.

Proof. Define $h(t) = \log\log\left(\frac{1}{t}\right)$, $g(t) = \frac{t}{\frac{\pi}{2} - \arctan(a)} h(t)$ consider two real numbers $0 < \beta < 1$ and $0 < \delta < 1$, take $t = \beta^{n-1}$ and $c = \frac{\frac{\pi}{2} + \theta}{t\sqrt{1+\delta}}$, then, Doob's maximal inequality shows that:

$$\begin{aligned}\mathbb{P}\left[\max_{0 \leq s \leq t}\{cZ_s\} > \sqrt{1+\delta}\, h(t)\right] &= \mathbb{P}\left[\max_{0 \leq s \leq t}\{\exp(cZ_s)\} > \exp\left(\sqrt{1+\delta}\, h(t)\right)\right] \\ &\leq \exp\left(-\sqrt{1+\delta}\, h(t)\right) E\left(\exp(cZ_t)\right)\end{aligned}$$

Using the upper bound (6.2) on the Laplace transform of Z, one gets:

$$\begin{aligned}\mathbb{P}\left[\max_{0 \leq s \leq t}\{cZ_s\} > \sqrt{1+\delta}\, h(t)\right] &\leq e^{-\sqrt{1+\delta} h(t)} \frac{1}{\rho t}\left[\frac{1}{\frac{\theta}{t} + \frac{\pi}{2t} - c} + \frac{1}{c - \frac{\theta}{t} + \frac{\pi}{2t}}\right] \\ &= e^{-\sqrt{1+\delta} h(t)}\left[\frac{1}{\theta + \frac{\pi}{2} - \frac{\frac{\pi}{2}+\theta}{\sqrt{1+\delta}}} + \frac{1}{\frac{\frac{\pi}{2}+\theta}{\sqrt{1+\delta}} - \theta + \frac{\pi}{2}}\right] \\ &= C(n-1)^{-\sqrt{1+\delta}}\end{aligned}$$

With C a constant independent of n. We obtain converging series, and we can a apply the Borel-Cantelli lemma:

$$\mathbb{P}\left[\max_{0 \leq s \leq \beta^{n-1}}\{Z_s\} \leq (1+\delta)\frac{1}{\frac{\pi}{2} - \arctan(a)} \beta^{n-1} h\left(\beta^{n-1}\right), n \to +\infty\right] = 1$$

$$Z_t \leq \max_{0 \leq s \leq \beta^{n-1}}\{Z_s\} \leq (1+\delta)\frac{1}{\beta} g(t) \frac{h\left(\beta^{n-1}\right)}{h\left(\beta^n\right)} \quad \text{a.s.} \tag{6.3}$$

Afterwards, taking $\beta \to 1$ and $\delta \to 0$ completes the proof. $\square$

The other inequality is slightly more complicated to derive. We begin with the preliminary lemma:

Lemma 6.8. $\mathbb{P}\left(Z_t - Z_h > z \middle| \mathcal{F}_h\right) \geq \frac{1}{4\rho}\frac{\exp\left(\left(\arctan(a)-\frac{\pi}{2}\right)\frac{Z}{t-h}\right)}{\frac{\pi}{2}-\arctan(a)}$ *for any* $t > h > 0$

Proof. First, remark that for any $t > 0$ and $0 < h < t$:

$$\begin{aligned} Z_t =& Z_h + \tilde{L}_{t-h} + \frac{a}{2}(\tilde{W}_{1,t-h}^2 + \tilde{W}_{2,t-h}^2) + \left(W_{1,h}\tilde{W}_{2,t-h} - W_{2,h}\tilde{W}_{1,t-h}\right) \\ &+ a\left(W_{1,h}\tilde{W}_{1,t-h} - W_{2,h}\tilde{W}_{2,t-h}\right) \\ Z_t =& Z_h + \tilde{Z}_{t-h} + \left\langle \gamma_h, \tilde{W}_{t-h}\right\rangle \end{aligned}$$

Denote $\tilde{W}_{t-h} = W_t - W_h$, which is independent of W_h. Define $\tilde{Z}_{t-h}, \tilde{L}_{t-h}$, as before, but with $\tilde{W}$ instead of W. Remark that γ_h is a vector which is a linear transform of W_h. Therefore, γ_h is independent of $\tilde{L}$,$\tilde{Z}$ and $\tilde{W}$. This independence is the key property to use the Borel-Cantelli lemma. Next, one shows that for any $z \in \mathbb{R}$:

$$\begin{aligned} \mathbb{P}\left(Z_t - Z_h > z \middle| \mathcal{F}_h\right) &= \mathbb{P}\left(\tilde{Z}_{t-h} + \left\langle \gamma_h, \tilde{W}_{t-h}\right\rangle > z \middle| \mathcal{F}_h\right) \\ &\geq \mathbb{P}\left(\tilde{Z}_{t-h} > z, \left\langle \gamma_h, \tilde{W}_{t-h}\right\rangle > 0 \middle| \mathcal{F}_h\right) \\ &\geq \frac{1}{2}\mathbb{P}\left(\tilde{Z}_{t-h} > z\right) \end{aligned}$$

The last inequality is due to the Brownian motion symmetry, and the independence property between γ_h and $\tilde{Z}$. Indeed taking $-\widetilde{W}$ instead of $\widetilde{W}$, does not change $\tilde{Z}$, and changes the sign of $\left\langle \gamma_h, \tilde{W}_{t-h}\right\rangle$. □

Now, we must prove that the Borel-Cantelli lemma holds for the sequence of events $\left\{Z_{t_n} - Z_{t_{n+1}} > z\right\}$ with t_n a decreasing sequence of times. We cannot use Borel-Cantelli directly as these events are not independent.

Lemma 6.9. *(Borel-Cantelli extension) Let* $t_n, n \in \mathbb{N}$ *be a decreasing sequence of positive numbers. And* z_n *a sequence of real numbers. If there exists a deterministic sequence* B_n *such that for any* $n \in \mathbb{N}$

$$\mathbb{P}\left(\left\{Z_{t_n} - Z_{t_{n+1}} > z_n\right\} \middle| \mathcal{F}_{t_{n+1}}\right) \geq B_n$$

almost surely, an if:

$$\sum_{n=0}^{+\infty} B_n = +\infty$$

Then, $Z_{t_n} - Z_{t_{n+1}} > z_n$ *infinitely often almost surely.*

Proof. The proof is much the same as the original Borel-Cantelli lemma's proof. Denoting A_n as the event $\left\{Z_{t_n} - Z_{t_{n+1}} \leq z_n\right\}$, and

$$B_n = \mathbb{P}\left(\left\{Z_{t_n} - Z_{t_{n+1}} > z_n\right\} \middle| \mathcal{F}_{t_{n+1}}\right)$$

One must prove that, for any $k \in \mathbb{N}$:

$$\mathbb{P}\left(\bigcap_{i=k}^{+\infty} A_i\right) = 0$$

As we have:

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i=k}^{+\infty} A_i\right) &= \mathbb{E}\left(\prod_{i=k}^{+\infty} \mathbb{1}_{A_i}\right) \\
&= \mathbb{E}\left[\mathbb{E}\left(\prod_{i=k}^{+\infty} \mathbb{1}_{A_i} \middle| \mathcal{F}_{t_{k+1}}\right)\right] \\
&= \mathbb{E}\left[\mathbb{E}\left(A_k | \mathcal{F}_{t_{k+1}}\right) \prod_{i=k+1}^{+\infty} \mathbb{1}_{A_i}\right] \text{ as } t_n \text{ is decreasing} \\
&\leq (1 - B_k)\mathbb{E}\left[\prod_{i=k+1}^{+\infty} \mathbb{1}_{A_i}\right]
\end{aligned}$$

Then, by a recursive argument, one obtains:

$$\mathbb{P}\left(\bigcap_{i=k}^{+\infty} A_i\right) \leq \prod_{i=k}^{+\infty} (1 - B_i)$$

At last, we use the inequality $1 - a \leq e^{-a}$ for any real number a. We obtain:

$$\mathbb{P}\left(\bigcap_{i=k}^{+\infty} A_i\right) \leq \exp - \sum_{i=k}^{+\infty} B_i \leq 0$$

As the series diverge by hypothesis. Then the proof is complete. □

With these considerations, one can prove the reverse limit inequality, that is to say:

Lemma 6.10. *The process Z is such that*

$$\limsup_{t \to 0^+} \frac{Z_t}{t \log\log \frac{1}{t}} \geq \frac{1}{\frac{\pi}{2} - \arctan(a)} > 0 \tag{6.4}$$

In the almost sure sense.

Proof. Denoting the event:

$$A_n = \left\{\left[Z\left(\beta^n\right) - Z\left(\beta^{n+1}\right)\right] > (1-\beta)^2 g\left(\beta^n\right)\right\}$$

Using lemma 6.8, this event occurs with probability:

$$\begin{aligned}
\mathbb{P}\left(A_n | \mathcal{F}_h\right) &\geq \frac{1}{4\rho} \frac{\exp\left(\left(\arctan(a) - \frac{\pi}{2}\right) \frac{(1-\beta)^2 g(\beta^n)}{\beta^n - \beta^{n+1}}\right)}{\frac{\pi}{2} - \arctan(a)} \\
&\geq C.n^{-(1-\beta)}
\end{aligned}$$

Which is the general term of some diverging series. Hence, the Borel-Cantelli lemma 6.9 leads to:

$$Z\left(\beta^n\right) > (1-\beta)^2 g\left(\beta^n\right) + Z\left(\beta^{n+1}\right)$$

infinitely often as $n \to +\infty$. Using the proof of lemma 6.7 considering inequality (6.3) leads to:

$$Z\left(\beta^{n+1}\right) < (1+\beta)\, g\left(\beta^{n+1}\right)$$

almost surely for n sufficiently large. Therefore one gets:

$$Z\left(\beta^n\right) > (1-\beta)^2 g\left(\beta^n\right) - (1+\beta)\, g\left(\beta^{n+1}\right) > (1-4\beta)\, g\left(\beta^n\right)$$

Infinitely often when $n \to +\infty$. So:

$$\limsup_{t\to 0} \frac{\int_0^t \left(\int_0^u \begin{pmatrix} a & 1 \\ -1 & a \end{pmatrix} dW_v\right)^T dW_u}{t \log\log\frac{1}{t}} \geq \frac{1-4\beta}{\frac{\pi}{2} - \arctan(a)}$$

For any $\beta > 0$. Taking $\beta \to 0$ finishes the proof. □

Remark 6.1. *At last, it is remarkable that the same kind of result could easily be derived in dimension n instead of 2. One would have to work a little bit on the matrices to obtain boundaries for the original n dimensional quadratic form. These boundaries would be quadratic forms expressed as block-diagonal matrices with 2×2 and scalar blocks. Then the same type of reasoning would give some estimates on the lower and upper limits, and would enable to treat the gamma constraints problem in n dimensions with non symmetric matrices.*

Chapter 2

Numerical approximation for superreplication problem under gamma constraints.

This is a joint work with Olivier Bokanowski[1], Stefania Maroso [2] and Haasna Zidani. [3]

In this chapter, we study the partial differential equation arising in chapter 1. We introduce a method to handle unbounded volatility controls that could be generalized to some Hamilton Jacobi Bellman equations. Many difficulties will come from the fact that the underlying process is a two dimensional degenerated diffusion. That is to say, the variance matrix is always of rank 1, and therefore standard results of finite difference methods do not apply. Hence, we use a generalized finite difference scheme that is consistent with degenerated diffusions. We will derive the theoretical properties of our algorithm in the viscosity solutions framework. Along this chapter, we will state the numerical scheme, prove its convergence, and finally give some numerical results.

Key words: Super-replication problem, viscosity solution, numerical approximation, monotone scheme, Howard algorithm

[1]Lab. Jacques Louis Lions, Univrsités Paris 6 & 7. E-mail: boka@math.jussieu.fr

[2]Projet SYDOCO, INRIA Rocquencourt. Also at Unité de Mathémaiques Appliquées, ENSTA Paris. E-mailStefania.Maroso@inria.fr

[3]Unité de Mathémaiques Appliquées, ENSTA Paris. Also at Projet SYDOCO, INRIA Rocquencourt. E-mail: Hasnaa.Zidani@ensta.fr

1 Introduction

In a financial market, consisting in a non-risky asset and some risky assets, people are interested to study the minimal initial capital needed in order to super-replicate a given contingent claim, under gamma constraints. Many authors have studied this problem in different cases and with different constraints: for example, see [29, 63], for problems in dimension 1, [20] for problems in dimension 2, and [64, 23] for problems in a general dimension d. In all these papers, authors characterize the super-replication price as the viscosity solution of an HJB-equation with terminal and boundary conditions. In a particular case, the dual formulation of the super-replication problem leads to a standard form of optimal stochastic control problem of [20].

In this paper we study numerically an HJB-equation coming from the super-replication problem in dimension 2. We discretize the HJB equation using the Generalized Finite Differences scheme [17, 18], then we study existence and uniqueness of the discrete solution. Finally we prove the convergence of the numerical solution to the viscosity solution. In particular, we are interested on the HJB equation which comes from the two dimensional dual problem introduced in [20]:

$$\vartheta(t,x,y) = \sup_{(\rho,\xi)\in\mathcal{U}} \mathbb{E}\left[g\left(X^{\rho,\xi}_{t,x,y}(T)\right)\right], \tag{1.1}$$

where (ρ,ξ) are valued in $[-1,1]\times(0,\infty)$, the process $(X^{\rho,\xi}_{t,x,y}, Y^{\rho,\xi}_{t,y})$ is a 2-dimensional positive process which evolves according to the stochastic dynamics (2.1), and g is a payoff function. The main difficulty of the above problem is due to the non-boundness of the control set, this fact implies that the Hamiltonian associated to (1.1) is not bounded, and numerical approximation for such a problem becomes more complicate.

In the literature, problems with unbounded control have been studied by many authors (for example, [1, 21]). In all these cases, the authors decide to truncate the set of controls to make it bounded. This truncation simplifies the numerical analysis of the problem. However, there is no theoretical result justifying this truncation.

In this paper we do not truncate the set of controls, because we find a particular form of our HJB equation which leads us to avoid the difficulty of unbounded control. In fact, our HJB equation can be reformulated in the following way

$$\Lambda^-(J(t,x,y,D\vartheta(t,x,y),D^2\vartheta(t,x,y))) = 0,$$

where J is a symetric matrix differential operator associated to the Hamiltonian, and where $\lambda^-(J)$ means the smallest eigenvalue of the matrix operator J. J does not depend on the control, but when we look for the first time at this equation, it seems that it is very difficult to treat. From standard computations on algebra, we rewrite the smallest eigenvalue as follows:

$$\Lambda^-(J) = \min_{\|\alpha\|=1} \alpha^T J\alpha,$$

where $\alpha\in\mathbb{R}^2$. Then we have transformed our problem into a bounded control problem, and now the numerical analysis is possible.

The structure of the paper is the following: in Section 2 we present the problem and the associated HJB-equation. We prove boundary conditions satisfied by the value function, then the existence, uniqueness and Lipschitz property of the viscosity solution. In Section

3 we consider the discretization of the HJB equation, and recall the main properties of the Generalized Finite Differences Scheme and we prove the consistency of this scheme. In section 4, we prove existence and uniqueness of a bounded discrete solution, and finally in Section 5 we prove the convergence of the numerical approximation.

2 Problem formulation and PDE

Let $(\Omega, \mathcal{F}_t, \mathbb{P})$ be a probability space, and $T > 0$ be a fixed finite time horizon. Let $\mathcal{U}$ denotes the set of all $\mathcal{F}_t$-measurable processes $(\rho, \zeta) := \{(\rho(t), \zeta(t)); 0 \leq t \leq T\}$ with values in $[-1, 1] \times \mathbb{R}_+$:

$$\mathcal{U} := \Big\{(\rho, \zeta) \text{ valued in } [-1, 1] \times (0, +\infty) \text{ and } \mathcal{F}_t\text{-measurable} \mid \int_0^T \zeta_t^2 \, dt < +\infty\Big\}.$$

For a given control process (ρ, ζ), and an initial data $(t, x, y) \in (0, T) \times \mathbb{R}^+ \times \mathbb{R}^+$, we consider the controlled 2-dimensional positive process $(X^{\rho,\zeta}_{t,x,y}, Y^{\rho,\zeta}_{t,y})$ evolving according to the stochastic dynamics:

$$dX^{\rho,\zeta}_{t,x,y}(s) = \sigma(s, Y^{\rho,\zeta}_{t,y}(s))X^{\rho,\zeta}_{t,x,y}(s)dW^1_s, \qquad s \in (t, T) \tag{2.1a}$$

$$dY^{\rho,\zeta}_{t,y}(s) = -\mu(s, Y^{\rho,\zeta}_{t,y}(s))ds + \zeta(s)Y^{\rho,\zeta}_{t,y}(s)dW^2_s, \quad s \in (t, T) \tag{2.1b}$$

$$\langle dW^1_s, dW^2_s \rangle = \rho(s), \qquad \text{a.e } s \in (t, T) \tag{2.1c}$$

$$X^{\rho,\zeta}_{t,x,y}(t) = x \ , \ Y^{\rho,\zeta}_{t,y}(t) = y, \tag{2.1d}$$

where W^1_s and W^2_s denote the standard Brownian motion defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The volatility σ and the cash flow μ satisfy the following assumptions:

(A1) $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}^+$ is a positive function, such that σ^2 is Lipschitz. For every $t \in [0, T]$, $\sigma(t, 0) = 0$ (typically $\sigma(t, y) = \sqrt{y}$).

(A2) $\mu : (0, T) \times \mathbb{R}^+ \to \mathbb{R}^+$ is a positive Lipschitz function, with $\mu(t, 0) = 0$ for every $t \in [0, T]$.

Assumptions (A1) and (A2) ensure that the stochastic dynamic system (2.1) has a unique strong solution.

The variables $X^{\rho,\zeta}_{t,x,y}$ and $Y^{\rho,\zeta}_{t,y}$ describe two different assets from a financial market. The first asset $X^{\rho,\zeta}_{t,x,y}$ is risky, while the second one $Y^{\rho,\zeta}_{t,y}$ distributes an instaneous cash flow $\mu(s, Y^{\rho,\zeta}_{t,y}(s))$, and its price is linked to the asset $X^{\rho,\zeta}_{t,x,y}$ by the means of volatility $\sigma(s, Y^{\rho,\zeta}_{t,y}(s))$.

Remark 2.1. *It is important to remark that the evolution of the variable* $Y^{\rho,\zeta}_{t,y}$ *does not depend on* $X^{\rho,\zeta}_{t,x,y}$.

Now consider a function $g : \mathbb{R}^+ \to \mathbb{R}$. Different assumptions will be made on g:

(A3) g is a bounded Lipschitz function. Let $M_0 > 0$ such that: $\|g\|_\infty \leq M_0$.

(A4) The function $f : z \to g\,(e^z)$ is Lipschitz continuous.

(A5) $g \in \mathcal{C}^2(\mathbb{R}^+ \to \mathbb{R})$. The functions $x \to xg'(x)$ and $x \to x^2 g''(x)$ are bounded.

Consider the following stochastic control problem $(\mathcal{P}_{t,x,y})$ with its associated value function ϑ defined by:

$$\vartheta(t,x,y) := \sup_{(\rho,\zeta)\in\mathcal{U}} \mathbb{E}\left[g\left(X^{\rho,\zeta}_{t,x,y}(T)\right)\right]. \tag{2.2}$$

Assumption (A3) leads us to obtain a bounded and Lipschitz value function ϑ of (2.2). Assumption (A4) will be usefull to prove some boundary conditions satisfied by ϑ (see section 2.1).

This control problem can be interpreted in [20] in the following sense: A trader wants to sell an European option of terminal payoff $g(X_T)$ without taking any risk. Hence we use a superreplication framework. The underlying X of the option is a risky asset, for axample a stock, an index or a mutual fund. Unfortunately, in several cases, the volatility σ of the underlying X exhibits large random changes across time. Therefore, the Black-Scholes model fails to capture the risks of the trader. One must then use a model that features stochastic volatility. It is known that in this framework, the superreplication problem has a trivial solution (see [29]). For example, if the volatility has no a priori bound, the superreplication price is the concave envelope of the payoff $g(X(T))$, and the hedging strategy is static. To obtain more accurate prices, we introduce another financial asset Y whose price is linked to the volatility of the underlying X. For example, we can consider a variance swap which continuously pays the instantaneous variance of X (hence $\mu(t,Y) = \sigma^2$). For the sake of simplicity we assume that the price of Y and the volatility of X are driven by a single common factor (hence $\sigma = \sigma(t,Y)$). If the parameters ζ and ρ of the dynamics of the price Y were known, and if there were no transaction costs for Y, the super-replication price would simply be $\mathbb{E}\left[g\left(X^{\rho,\zeta}_{t,x,y}(T)\right)\right]$. But we face two problems:

- The parameters (ζ,ρ) of the dynamics of Y are likely to be random and difficult to measure. As there is no a priori bound to these parameters, the super-replication price is given be the supremum of $\mathbb{E}\left[g\left(X^{\rho,\zeta}_{t,x,y}(T)\right)\right]$ over all adapted processes ζ,ρ (see [36]).

- The asset Y is likely to introduce transaction costs, and hence the trader cannot buy and sell an infinite amount of asset Y during the period $[0,T]$. It is proved in [20] that the super-replication price of $g(X(T))$ under the constraint of a finite amount of transactions involving Y during $[0,T]$ is given by the value function of problem (2.2). See also [63, 64] for a similar approach.

Denote by $\mathcal{M}_2$ the set of symmetric 2×2 matrices. The Hamiltonian function is defined by: for $t \in (0,T)$, $x,y \in \mathbb{R}^+$, $p = (p_1,p_2)^\mathsf{T} \in \mathbb{R}^2$, and $Q \in \mathcal{M}_2$:

$$H(t,x,y,p,Q) := \inf_{(\zeta,\rho)\in\mathbb{R}_+\times[-1,1]} \left\{\mu(t,y)p_2 - \frac{1}{2}\mathrm{tr}\left(a(t,x,y,\zeta,\rho)\cdot Q\right)\right\}, \tag{2.3}$$

and the covariance matrix a is given by:

$$a(t,x,y,\zeta,\rho) := \begin{pmatrix} \sigma^2(t,y)x^2 & \rho\zeta\sigma(t,y)x \\ \rho\zeta\sigma(t,y)x & \zeta^2 \end{pmatrix}.$$

Now we look for a characterization of ϑ as a viscosity solution of an HJB equation. In a formal way, we can check that ϑ satisfies the following PDE:

$$-\frac{\partial \vartheta}{\partial t} + H\left(t,x,y,D\vartheta,D^2\vartheta\right) = 0 \quad (t,x,y) \in (0,T)\times(0,+\infty)\times(0,+\infty). \tag{2.4}$$

We will prove in Theorem 2.3 that the precise HJB equation satisfied by ϑ in the viscosity sens is

$$\Lambda^- \begin{pmatrix} -\frac{\partial\vartheta}{\partial t} + \mu(t,y)\frac{\partial\vartheta}{\partial y} - \frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} \\ -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2} \end{pmatrix} = 0, \tag{2.5}$$

where $\Lambda^-(A)$ denotes the smallest eigenvalue of a given symmetric matrix A. We first prove that ϑ is a discontinuous viscosity solution of (2.5). We will see later on that, under (A1), ϑ is continuous thanks to a comparison principle, and even Lipschitz continuous when assumptions (A3)-(A5) hold.

First, it is easy to see that the infimum in (2.3) can only be achieved for $\rho = \pm 1$. Hence denoting ζ as $\rho\zeta$, one can see that the Hamiltonian can be rewritten as:

$$H\left(t,x,y,p,Q\right) = \inf_{\zeta\in\mathbb{R}} \left\{ \mu(t,y)p_2 - \frac{1}{2}\mathrm{tr}\left(a(t,x,y,\zeta)\cdot Q\right) \right\}, \tag{2.6}$$

where, this time, there is only one control variable ζ taking values on the whole real line, and the covariance matrix a is defined by:

$$a(t,x,y,\zeta) = \begin{pmatrix} \sigma^2(t,y)x^2 & \zeta\sigma(t,y)x \\ \zeta\sigma(t,y)x & \zeta^2 \end{pmatrix}.$$

By elementary techniques, the minimization over ζ, in (2.6) gives:

$$H(t,x,y,p,Q) = -\infty \qquad \text{if } Q_{22} > 0, \tag{2.7a}$$

$$\text{or } \ Q_{22} = 0 \text{ and } \sigma(t,y)xQ_{12} \neq 0, \tag{2.7b}$$

$$H(t,x,y,p,Q) \in \mathbb{R}, \qquad \text{otherwise.} \tag{2.7c}$$

Remark 2.2. *For this particular problem, it is not possible to find a continuous function* $G : [0,T]\times\mathbb{R}^2\times\mathbb{R}^2_+\times\mathcal{M}_2 \to \mathbb{R}$ *such that*

$$H(t,x,y,p,Q) > -\infty \Leftrightarrow G(t,x,y,p,Q) \geq 0.$$

Hence we can not use arguments introduced in [58] to obtain the HJB equation (2.5).

For $t \in (0,T)$, $x,y \in \mathbb{R}^+$, $r \in \mathbb{R}$, $p = (p_1,p_2)^T \in \mathbb{R}^2$ and $Q \in \mathcal{M}_2$, introduce the notation:

$$J(t,x,y,r,p,Q) = \begin{pmatrix} -r + \mu(t,y)p_2 - \frac{1}{2}\sigma^2(t,y)x^2Q_{11} & -\frac{1}{2}\sigma(t,y)xQ_{12} \\ -\frac{1}{2}\sigma(t,y)xQ_{12} & -\frac{1}{2}Q_{22} \end{pmatrix}.$$

With straightforward computations we obtain the following result.

Lemma 2.11. *For $t \in (0,T)$, $x, y \in \mathbb{R}^+$, $r \in \mathbb{R}$, $p = (p_1, p_2)^T \in \mathbb{R}^2$ and $Q \in \mathcal{M}_2$, the following assertions hold:*

(i) $-r + H(t,x,y,p,Q) \geq 0 \Leftrightarrow \Lambda^-(J(t,x,y,r,p,Q)) \geq 0.$

(ii) $-r + H(t,x,y,p,Q) \geq 0 \Rightarrow -Q_{22} \geq 0.$

(iii) $-r + H(t,x,y,p,Q) = 0 \Rightarrow \Lambda^-(J(t,x,y,r,p,Q)) = 0.$

(iv) $\Lambda^-(J(t,x,y,r,p,Q)) > 0 \Rightarrow -r + H(t,x,y,p,Q) > 0.$

Now, for a function $u : [0,T] \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$, we define the upper (resp. lower) semicontinuous envelope u^* (resp. u_*) of u by : for $t \in [0,T), x, y \in (0,+\infty)$,

$$u^*(t,x,y) = \limsup_{\substack{(s,w,z)\to(t,x,y) \\ s\geq 0, w,z\in(0,+\infty)}} u(s,w,z),$$
$$u_*(t,x,y) = \liminf_{\substack{(s,w,z)\to(t,x,y) \\ s\geq 0, w,z\in(0,+\infty)}} u(s,w,z).$$

With these definitions, we can give the sens of viscosity solution of (2.5), according to [5, 6, 27].

Definition 2.3. *(i) u is a discontinuous viscosity subsolution of* (2.4) *iff for any $(\hat{t},\hat{x},\hat{y}) \in [0,T)\times(0,+\infty)^2$, and any $\phi \in C^2\left([0,T)\times(0,+\infty)^2\right)$, such that $(\hat{t},\hat{x},\hat{y})$ is a local maximum of $u^* - \phi$:*

$$\Lambda^-(J(\hat{t},\hat{x},\hat{y}), \partial_t\phi(\hat{t},\hat{x},\hat{y}), D\phi(\hat{t},\hat{x},\hat{y}), D^2\phi(\hat{t},\hat{x},\hat{y}))) \leq 0.$$

(ii) u is a discontinuous viscosity super-solution of (2.4) *iff for any $(\hat{t},\hat{x},\hat{y}) \in [0,T) \times (0,+\infty)^2$, and any $\phi \in C^2\left([0,T)\times(0,+\infty)^2\right)$, such that $(\hat{t},\hat{x},\hat{y})$ is a local minimum of $u_* - \phi$:*

$$\Lambda^-(J(\hat{t},\hat{x},\hat{y}), \partial_t\phi(\hat{t},\hat{x},\hat{y}), D\phi(\hat{t},\hat{x},\hat{y}), D^2\phi(\hat{t},\hat{x},\hat{y}))) \geq 0.$$

(iii) u is a discontinuous viscosity solution of (2.4) *iff it is both sub and a super solution.*

Theorem 2.3. *Under assumptions (A1)-(A2), the value function ϑ is a viscosity discontinuous solution of* (2.5)*:*

$$\Lambda^-\begin{pmatrix} -\frac{\partial\vartheta}{\partial t} + \mu(t,y)\frac{\partial\vartheta}{\partial y} - \frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} \\ -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2} \end{pmatrix} = 0.$$

Moreover ϑ is a discontinuous viscosity super-solution of

$$-\frac{\partial^2\vartheta}{\partial y^2} \geq 0. \tag{2.8}$$

Proof. The proof is splitted on two parts: the super-solution property and the sub-solution property.

(a) Super-solution property. By a classical application of the Dynamic Programming Principle, as done in [51], we obtain that $\vartheta(t,x,y)$ is a viscosity super-solution of

$$-\frac{\partial\vartheta}{\partial t} + H(t,x,y,D\vartheta,D^2\vartheta) \geq 0.$$

Then, Lemma 2.11(i) implies that also

$$\Lambda^-(J(t,x,y,\partial_t\vartheta,D\vartheta,D^2\vartheta)) \geq 0,$$

and then ϑ is also a viscosity super-solution of (2.5).

Moreover, this last inequality implies that $-\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2} \geq 0$, and hence (2.8) is verified. **(b) Sub-solution property.** Let φ be a smooth function, and let $(\bar{t},\bar{x},\bar{y})$ be a strict maximizer of $\vartheta^* - \varphi$, such that

$$0 = (\vartheta^* - \varphi)(\bar{t},\bar{x},\bar{y}).$$

Suppose that $(\bar{t},\bar{x},\bar{y})$ belongs to the set $\mathcal{M}(\varphi)$ defined by:

$$\mathcal{M}(\varphi) = \{(t,x,y) \in [0,T)\times(0,+\infty)^2 : \Lambda^-(J(t,x,y,\partial_t\varphi(t,x,y),D\varphi(t,x,y),D^2\varphi(t,x,y))) > 0\}$$

Since $\mathcal{M}(\varphi)$ is an open set, then there exists $\eta > 0$ such that

$$[0 \wedge (\bar{t}-\eta),\bar{t}+\eta] \times \overline{B}_\eta(\bar{x},\bar{y}) \subset \mathcal{M}(\varphi),$$

where $\overline{B}_\eta(\bar{x},\bar{y})$ denotes the closed ball centered in $(\bar{x},\bar{y})$ and with radius η. From Lemma 2.11(iii), if $(t,x,y) \in \mathcal{M}(\varphi)$, then

$$-\frac{\partial\varphi}{\partial t}(t,x,y) + H(t,x,y,D\varphi(t,x,y),D^2\varphi(t,x,y)) > 0.$$

Using the Dynamic Programming Principle and the same arguments that in [58, Lemma 3.1], we get that:

$$\sup_{\partial_p([0\wedge(\bar{t}-\eta),\bar{t}+\eta]\times\overline{B}_\eta(\bar{x},\bar{y}))} (\vartheta - \varphi) = \max_{[0\wedge(\bar{t}-\eta),\bar{t}+\eta]\times\overline{B}_\eta(\bar{x},\bar{y})} (\vartheta^* - \varphi), \tag{2.9}$$

where $\partial_p([t_1,t_2]\times\overline{B}_\eta(\bar{x},\bar{y}))$ is the forward parabolic boundary of $[t_1,t_2]\times\bar{B}_\eta(\bar{x},\bar{y})$, i.e. $\partial_p([t_1,t_2]\times\overline{B}_\eta(\bar{x},\bar{y})) = [t_1,t_2]\times\partial\overline{B}_\eta(\bar{x},\bar{y})\cup\{t_2\}\times\overline{B}_\eta(\bar{x},\bar{y})$. However, since $(\bar{t},\bar{x},\bar{y})$ is a strict maximizer of $\vartheta^* - \varphi$, equality (2.9) leads to a contradiction. Therefore, $(\bar{t},\bar{x},\bar{y}) \notin \mathcal{M}(\varphi)$, and the result follows. □

In our paper, we are interested by the numerical computation of the value function ϑ. Although equation (2.5) has a rigorous meaning, the formulation with the smallest eigenvalue makes difficult to deal with its numerical discretization. Of course, one can be tempted to modify the hamiltonian in the following way: for $\zeta_{\max} > 0$,

$$H(t,x,y,p,Q) \cong \min_{\zeta\in[-\zeta_{\max},\zeta_{\max}]} \left\{\mu(t,y)p_2 - \frac{1}{2}\mathrm{tr}(a(t,x,y,\zeta)\cdot Q)\right\}.$$

However, the choice of $\zeta_{\max}$, guaranteeing a good approximation of H, does not appear obvious to us. To avoid these difficulties, we first give an equivalent HJB equation satisfied by ϑ and which is formulated with bounded controls. More precisely, we have:

Corollary 2.1. *Under assumptions (A1)-(A3), the value function ϑ is a viscosity solution of the HJB equation:*

$$\inf_{\alpha_1^2+\alpha_2^2=1}\left\{\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}^T \begin{pmatrix} -\frac{\partial\vartheta}{\partial t} + \mu(t,y)\frac{\partial\vartheta}{\partial y} - \frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} \\ -\frac{1}{2}\sigma(t,y)x\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\frac{\partial^2\vartheta}{\partial y^2}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}\right\} = 0. \tag{2.10}$$

Remark 2.4. *Equation* (2.5) *can be reformulated as follows,*

$$\Lambda^-\begin{pmatrix} -\frac{\partial\vartheta}{\partial t}+\mu(t,y)\frac{\partial\vartheta}{\partial y}-\frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2\vartheta}{\partial x^2} & -\frac{1}{2}\sigma(t,y)x\eta(t,y)\frac{\partial^2\vartheta}{\partial x\partial y}\\ -\frac{1}{2}\sigma(t,y)x\eta(t,y)\frac{\partial^2\vartheta}{\partial x\partial y} & -\frac{1}{2}\eta^2(t,y)\frac{\partial^2\vartheta}{\partial y^2}\end{pmatrix}=0, \tag{2.11}$$

where $\eta(t,y)$ is any strictly positive function. l It is easy to see that changing the positive function $\eta(t,y)$ into another positive function, does not change the sign of the operator in (2.5)*, for fixed $(t,x,y,D\vartheta,D^2\vartheta)$.*

In particular, when we will deal with the discretization of (2.10)*, we will use* $\eta(t,y)=\min(1,y)$.

2.1 Boundary conditions. Uniqueness result

Unlike in most similar parabolic problems, here we do not only need a terminal condition to obtain the uniqueness, but also a border conditions when y tends to zero. Another boundary condition is hidden by the fact that we only consider bounded solutions, which is, intuitively, equivalent to Neumann conditions near infinity.

Lemma 2.12. *Under assumptions (A1)-(A3), the value function ϑ is bounded and satisfies the following conditions on the boundaries $x=0$ and $y=0$:*

$$\lim_{(t',x',y')\to(t,x,0)}\vartheta(t',x',y')=\vartheta(t,x,0)=g(x),\forall\,(t,x)\in[0,T]\times\mathbb{R}_+^* \tag{2.12a}$$

$$\lim_{(t',x',y')\to(t,0,y)}\vartheta(t',x',y')=\vartheta(t,0,y)=g(0),\forall\,(t,y)\in[0,T]\times\mathbb{R}_+^* \tag{2.12b}$$

and the terminal condition of the equation for $t=T$ is:

$$\lim_{(t',x',y')\to(T,x,y)}\vartheta(t',x',y')=\vartheta(T,x,y)=g(x)\ \textit{for all}\ (x,y)\in(\mathbb{R}_+^*)^2. \tag{2.12c}$$

Proof. The statements (2.12a)-(2.12c) are proved in lemma 5.6 in [20]. The proof is based on the assumptions (A1) and (A2) of σ and μ, and on the continuity and boundedness of g (see (A3)).

Now to prove statement (2.12b), we first give a representation of $\vartheta(t,x,y)$ using Doleans integral. Indeed, for every (t,x,y), we have:

$$X_{t,x,y}^{\rho,\zeta}=xZ_y^{\zeta,\rho},\quad\text{where}\quad Z_y^{\zeta,\rho}:=e^{\int_t^T\sigma(s,Y_{t,y}^{\rho,\zeta}(s)dW_s^1+\frac{1}{2}\int_t^T(\sigma(s,Y_{t,y}^{\rho,\zeta}(s)))^2ds}.$$

Therefore,

$$\vartheta(t,x,y)=\mathbb{E}\Big[g(X_{t,x,y}^{\rho,\zeta})(T)\Big]=\mathbb{E}\Big[g\Big(xZ_y^{\zeta,\rho},\Big)\Big]. \tag{2.13}$$

We conclude that statements (2.12b) holds. □

We recall here the uniqueness result, proved in Lemma 4.3, Proposition 4.4, and Proposition 4.6 of [20].

Theorem 2.5. *(Proposition 4.4 of [20]) Assume (A1)-(A3). Suppose that u is an upper semi-continuous viscosity sub-solution of* (2.5) *bounded from above, and w a lower semi-continuous viscosity super-solution of* (2.5) *bounded from below. If, furthermore,*

$$\begin{aligned} u(T,x,y) &\le g(x) \le w(T,x,y),\\ u(t,x,0) &\le g(x) \le w(t,x,0), \end{aligned} \tag{2.14}$$

then $u(t,x,y) \le w(t,x,y)$, for all $(t,x,y) \in [0,T] \times \mathbb{R}^2_+$. In particular, the solution of (2.5) *in the viscosity sense with boundary conditions* (2.12a) *and* (2.12c) *is unique.*

We recall here the main ideas of the proof.

Proof. Suppose that u and w are respectively sub- and super-solution of (2.5), and that they both satisfy the limit conditions (2.12a), (2.12b) and (2.12c). A classical argument (see [7]) to prove uniqueness for equation as (2.5), consists in building a strict viscosity super-solution of (2.5) w_ε, depending on the super-solution and on a parameter ε. Moreover w_ε must to be such that, when the parameter ε goes to zero, w_ε tends to w. Then with classical arguments [27], a comparison principle between the strict super-solution and the sub-solution can be obtained, and sending ε to zero we have the desired comparison principle.

In our particular case, for any $\varepsilon > 0$, we build

$$w_\varepsilon = w + \varepsilon((T-t) + \ln(1+y)).$$

From Lemma 4.3 of [20], w_ε is a strict viscosity super-solution of (2.5), bounded from below and such that conditions (2.14) are satisfied. Then we can apply Proposition 4.6 of [20] which is a comparison principle between a strict viscosity super-solution and a viscosity sub-solution, and we obtain

$$w_\varepsilon \ge u,$$

for all $(t,x,y) \in [0,T] \times \mathbb{R}^2_+$. Sending ε to zero, we have the result. $\square$

Since the boundedness property of ϑ would be tricky to manipulate numerically, in the following proposition we give some growth properties of the value function which are a sort of Neumann conditions at infinity. These conditions will guide us toward an implementable scheme.

Proposition 2.15. *Assume that (A1)-(A4) are satisfied. Then the following holds:*
(i) For any $a > 0$, the function:

$$h^1_{t,y} : x \to \vartheta(t, x+a, y) - \vartheta(t,x,y)$$

converges to zero, uniformly in (t,y), when $x \to +\infty$.
(ii) The function;

$$h^2_{t,x} : y \to \vartheta(t, x, y+a) - \vartheta(t,x,y)$$

converges to zero, uniformly in (t,x), when $y \to +\infty$.

Proof. (i) Let $(t,x,y) \in (0,T) \times \mathbb{R}^+ \times \mathbb{R}^+$. As in (2.13), we have:

$$\vartheta(t,x,y) = \sup_{\zeta,\rho} \mathbb{E}\left[g\left(X^{\rho;\zeta}_{t,x,y}(T)\right)\right] = \sup_{\zeta,\rho} \mathbb{E}\left[g\left(xZ^{\zeta,\rho}_y\right)\right]. \tag{2.15}$$

By assumption (A3), the function $f : z \to g(e^z)$ is Lipschitz continuous on $\mathbb{R}$. Then, for $x' \in \mathbb{R}^+$, we get:

$$\begin{aligned}
\vartheta(t,x,y) - \vartheta(t,x',y) &= \sup_{\zeta,\rho} \mathbb{E}\left(g\left(xZ^{\zeta,\rho}_y\right)\right) - \sup_{\zeta,\rho} \mathbb{E}\left(g\left(x'Z^{\zeta,\rho}_y\right)\right) \\
&\le \sup_{\zeta,\rho}\left\{\mathbb{E}\left(g\left(xZ^{\zeta,\rho}_y\right)\right) - \mathbb{E}\left(g\left(x'Z^{\zeta,\rho}_y\right)\right)\right\} \\
&\le \sup_{\zeta,\rho}\left\{\mathbb{E}\left[f\left(\ln(x) + \ln\left(Z^{\zeta,\rho}_y\right)\right) - f\left(\ln(x') + \ln\left(Z^{\zeta,\rho}_y\right)\right)\right]\right\},
\end{aligned}$$

and using the Lipschitz property of f, it yields to:

$$\vartheta(t,x,y) - \vartheta(t,x',y) \leq K \left|\ln(x) - \ln(x')\right|.$$

Therefore we get that

$$\left|h^1_{t,y}(x)\right| \leq K \left|\ln\left(\frac{x+a}{x}\right)\right| \to 0 \text{ as } x \to +\infty \text{ uniformly in } (t,y). \tag{2.16}$$

To prove assertion (ii), using (2.8), we see that ϑ is a supersolution of

$$-\frac{\partial^2 v}{\partial y^2} = 0.$$

Then, from [29], we deduce that the function ϑ is concave w.r.t. y. That is, for each (t,x), $\vartheta(t,x,\cdot)$ is a concave function. Moreover, from (A3), ϑ is bounded and $\|\vartheta\|_\infty \leq M_0$ (where the constant $M_0 > 0$ is the same as in (A3)). Therefore, for any λ, the function

$$h^2_{t,x} : y \to \vartheta(t,x,y+\lambda) - \vartheta(t,x,y)$$

is decreasing. Considering that $\vartheta(t,x,n\lambda+y_0) = \vartheta(t,x,y_0) + \sum_{i=1}^n h^2_{t,x}(i\lambda+y_0)$. Hence, it follows that:

$$\vartheta(t,x,n\lambda+y_0) \geq \vartheta(t,x,y_0) + \sum_{i=1}^{n} h^2_{t,x}(n\lambda+y_0)$$

which gives:

$$h^2_{t,x}(n\lambda+y_0) \leq \frac{2M}{n}$$

and we get convergence of $h^2_{t,x}(y)$ to 0, which is uniform in (t,x). □

2.2 Lipschitz property

Here we establish the Lipschitz property of the value function ϑ.

Proposition 2.16. *Under assumptions (A1)-(A4), we have:*

(i) The value function ϑ is Lipschitz w.r.t. x.

(ii) ϑ is Lipschitz w.r.t. y.

Proof. (i)As in the proof of proposition 2.15, we consider the representation of ϑ using Doleans exponential:

$$\vartheta(t,x,y) = \sup_{\zeta,\rho} \mathbb{E}\left(g(X^{\zeta,\rho}_{t,x,y})\right) = \sup_{\zeta,\rho} \mathbb{E}\left[g\left(xZ^{\zeta,\rho}_y\right)\right] \qquad \forall t \in (0,T), x,y \in \mathbb{R}^+, \tag{2.17}$$

where $Z^{\zeta,\rho}_y = e^{\int_t^T \sigma(s,Y^{\rho,\zeta}_{t,y}(s)dW^1_s + \frac{1}{2}\int_t^T(\sigma(s,Y^{\rho,\zeta}_{t,y}(s)))^2 ds}$.

Then, for $t \in (0,T), x, x', y \in \mathbb{R}^+$ we have:

$$\left|\vartheta(t,x,y) - \vartheta(t,x',y)\right| \leq \sup_{\zeta,\rho} \mathbb{E}\left[g\left(xZ^{\zeta,\rho}_y\right) - g\left(xZ^{\zeta,\rho}_y\right)\right].$$

As g is Lipschitz of constant K, we get:

$$\left|\vartheta(t,x,y)-\vartheta(t,x',y)\right| \le \sup_{\zeta,\rho}\mathbb{E}\left|K(x-x')Z_y^{\zeta,\rho}\right| \le K|x-x'|\sup_{\zeta,\rho}\mathbb{E}\left(Z_y^{\zeta,\rho}\right).$$

Therefore, using the fact that the Doleans exponential is a positive local martingale, and hence a super-martingale, which implies that for any control $(\zeta,\rho)\in\mathcal{U}$:

$$\mathbb{E}\left(e^{\int_t^T \sigma_u^{\zeta,\rho}dW_u+\frac{1}{2}\int_t^T(\sigma_u)^{\zeta,\rho})^2du}\right) \le 1,$$

and then taking the supremum leads to:

$$\left|\vartheta(t,x,y)-\vartheta(t,x',y)\right| \le K|x-x'|$$

Which proves that ϑ is Lipschitz w.r.t. x with the same constant as g.

(ii) Now we treat the Lipschitz property of ϑ w.r.t. y.

First,we recall that ϑ is concave w.r.t. y. Furthermore, as g is bounded, we immediately get that ϑ shares the same bound. Hence, it is sufficient to prove that ϑ is Lipschitz near the boundary $y=0$.

Recall that by (2.12a), we know that $\vartheta(t,x,0)=g(x)$ for all $(t,x)\in(0,T)\times(0,+\infty)$.

Let $(t,x,y)\in[0,T]\times(0,+\infty)^2$, with $y>0$. For any control $(\zeta,\rho)\in\mathcal{U}$, we have:

$$Y_{t,y}^{\rho,\zeta}(s) = y+\int_t^s -\mu(\tau,Y_{t,y}^{\rho,\zeta}(\tau))d\tau+\int_t^s \zeta(\tau)Y_{t,y}^{\rho,\zeta}\tau\, dW_\tau^2.$$

Furthermore, by a comparison argument for SDEs, we get, for any $\tau\in[t,T]$:

$$Y_{t,y}^{\rho,\zeta}(\tau)\ge 0.$$

Using the positivity of μ, we get:

$$0\le Y_{t,y}^{\rho,\zeta}(s)\le y+\int_t^s Y_{t,y}^{\rho,\zeta}(\tau)dW_\tau^2.$$

Hence, the quantity above is a super-martingale and we get:

$$\mathbb{E}\left[Y_{t,y}^{\rho,\zeta}(s)\right]\le y. \tag{2.18}$$

Now, applying Itô's formula on $g(X_{t,x,y}^{\rho,\zeta})$:

$$\begin{aligned}
g(X_{t,x,y}^{\rho,\zeta}(s)) &= g(x)+\int_t^s g'(X_{t,x,y}^{\rho,\zeta}(\tau))dX_{t,x,y}^{\rho,\zeta}(\tau)+\\
&\quad \frac{1}{2}\int_t^s g''(X_{t,x,y}^{\rho,\zeta}(\tau))\left\langle dX_{t,x,y}^{\rho,\zeta}(\tau),dX_{t,x,y}^{\rho,\zeta}(\tau)\right\rangle\\
&= g(x)+\int_t^s g'(X_{t,x,y}^{\rho,\zeta}(\tau))dX_{t,x,y}^{\rho,\zeta}(\tau)+\\
&\quad \frac{1}{2}\int_t^s \left(X_{t,x,y}^{\rho,\zeta}(\tau)\right)^2 g''(X_{t,x,y}^{\rho,\zeta}(\tau))\sigma^2(Y_{t,y}^{\rho,\zeta}(\tau))d\tau.
\end{aligned}$$

Since $X^{\rho,\zeta}_{t,x,y}$ is a locale martingale, there exists a sequence $(s_n)_n$, with $s_n \to \infty$ such that:

$$\mathbb{E}\Big(\int_t^{s_n\wedge T} g'(X^{\rho,\zeta}_{t,x,y}(u))dX^{\rho,\zeta}_{t,x,y}(u)\Big) = 0.$$

Using (2.18), the Lipschitz property of σ^2, and the boundedness of $x \longmapsto x^2 g''(x)$, it yields: there exists a constant $C > 0$, such that:

$$\Big|\mathbb{E}\left(g(X^{\rho,\zeta}_{t,x,y}(s_n\wedge T)) - g(x)\right)\Big| \leq \int_t^{s_n\wedge T} Cy\, d\tau$$

Finally, as g is bounded, we conclude with Fatou's lemma that:

$$\mathbb{E}\left(g(X^{\rho,\zeta}_{t,x,y}(T)) - g(x)\right) \leq C(T-t)y,$$

and since the constant C is independent of ρ, ζ, we obtain:

$$\begin{aligned}|\vartheta(t,x,y) - \vartheta(t,x,0)| &\leq \sup_{(\rho,\zeta)\in\mathcal{U}} \left\{|\mathbb{E}\left(g(X^{\rho,\zeta}_{t,x,y}(T)) - g(x)\right)|\right\}\\ &\leq CTy.\end{aligned}$$

Hence, as ϑ is concave w.r.t. y and bounded, it is Lipschitz with respect to y. □

3 Approximation Scheme

Now we want to approximate the (unique) bounded solution of the following Hamilton-Jacobi-Bellman equation:

$$\min_{\alpha_1^2+\alpha_2^2=1}\left\{ -\alpha_1^2\frac{\partial\vartheta}{\partial t}(t,x,y) + \mu(t,y)\alpha_1^2\frac{\partial\vartheta}{\partial y}(t,x,y) - \frac{1}{2}\mathrm{tr}[a(\alpha_1,\alpha_2,t,x,y)D^2\vartheta(t,x,y)]\right\} = 0, \tag{3.1}$$

with boundary conditions (2.12a), (2.12b), (2.12c), where μ is a positive Lipschitz function, and the diffusion matrix a is defined as follows:

$$\begin{aligned} a(\alpha_1,\alpha_2,t,x,y) &:= \begin{pmatrix} \alpha_1^2\sigma^2(t,y)x^2 & \alpha_1\alpha_2\sigma(t,y)\eta(t,y)x \\ \alpha_1\alpha_2\eta(t,y)\sigma(t,y)x & \eta^2(t,y)\alpha_2^2 \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1\sigma(t,y)x \\ \eta(t,y)\alpha_2 \end{pmatrix}\begin{pmatrix} \alpha_1\sigma(t,y)x \\ \alpha_2\eta(t,y) \end{pmatrix}^{\mathsf{T}}, \end{aligned} \tag{3.2}$$

where $\eta(t,y) = \min(1,y)$, in agreement with (2.11). From now on we will write only a instead of $a(\alpha_1,\alpha_2,t,x,y)$, and μ instead of $\mu(t,y)$, we omit all the dependences.

We can easily see that a is not a dominant diagonal matrix[4], in fact we can not ensure that

$$\alpha_2\eta \geq \alpha_1\sigma x, \quad \forall\,(t,x,y)\in[0,T)\times[0,+\infty)^2, \text{ and } \forall\,\alpha_1^2+\alpha_2^2=1.$$

[4]We recall that a matrix X of dimension $N\times N$ is diagonal dominant if

$$X_{ii} \geq \sum_{i\neq j}|X_{ij}|, \quad \forall\, i=1,\dots,N.$$

This fact implies that we can not choose the Classical Finite Differences scheme to approximate equation (3.1), we shall use the Generalized Finite Differences scheme introduced in [17].

Consider a regular grid G_h of discretization of $\mathbb{R}^2_+$, with discretization steps $h = (h_1, h_2)$:

$$G_h := \Big\{(x_i, y_j),\ x_i := ih_1,\ y_j := jh_2,\ i, j \in \mathbb{N} \times \mathbb{N}\Big\},$$

and consider a discretization time step Δt. On the grid G_h, the derivative on time is approximated by an implicit Euler scheme, and for the first derivative in y we use a finite difference approximation. The main idea of the Generalized Finite Differences scheme is to approximate the diffusion term $a \cdot D^2\phi$ by a linear combination of elementary diffusions pointing towards grid points. More precisely, for $\xi = (\xi_1, \xi_2) \in \mathbb{Z}^2$, associate the second order finite difference operator (for $x, y \in \mathbb{R}$):

$$\Delta_\xi \phi(t, x, y) = \phi(t, x + \xi_1 h_1, y + \xi_2 h_2) + \phi(t, x - \xi_1 h_1, y - \xi_2 h_2) - 2\phi(t, x, y),$$

where Δ_ξ is an elementary diffusion in the direction ξ. By a Taylor expansion, we know that

$$\Delta_\xi \phi(t, x, y) = \sum_{i,j=1}^{2} h_i h_j \xi_i \xi_j \phi_{x_i x_j} + o(||h^2||),$$

where $x_1 = x$ and $x_2 = y$.

Following ([17, 18]), we introduce a set $\mathcal{S} \subseteq \mathbb{Z}^2 \setminus 0$, which contains $\{e_1, e_2\}$. We will specify later how we choose this set. We approximate the second order term $a \cdot D^2\phi$ by a linear combination of elementary diffusions along ξ, with $\xi \in \mathcal{S}$:

$$a \cdot D^2\phi \cong \sum_{\xi \in \mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \Delta_\xi \phi,$$

where the $\gamma_\xi^{\alpha_1,\alpha_2}$ are coefficients which will be specified later.

For a given set $\mathcal{S}$, the scheme takes the following form:

$$v_h(T, x, y) = g(x) = v_h(t, x, 0), \quad v_h(t, 0, y) = g(0), \tag{3.3}$$

$$\min_{\alpha_1^2+\alpha_2^2=1} \{-\alpha_1^2 \delta_t v_h(t, x, y) - \alpha_1^2 \mu \delta_y v_h(t, x, y) - \frac{1}{2} \sum_{\xi \in \mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \Delta_\xi v_h(t, x, y)\} = 0, \tag{3.4}$$

for $t < T - \Delta t$, with

$$\begin{aligned} \delta_t v_h(t, x, y) &= \frac{v_h(t + \Delta t, x, y) - v_h(t, x, y)}{\Delta t}, \\ \delta_y v_h(t, x, y) &= \frac{v_h(t, x, y - h_2) - v_h(t, x, y)}{h_2}. \end{aligned}$$

It is shown in [17, 18] that the above scheme is consistent if we choose a set $\mathcal{S}$ and variables $\gamma_\xi^{\alpha_1,\alpha_2}$ such that: for all α_1, α_2, t, x, y

$$\begin{aligned} &\gamma_\xi^{\alpha_1,\alpha_2} \geq 0, \ \forall\ \xi \in \mathcal{S}, \\ &\sum_{\xi \in \mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \xi \xi^\mathsf{T} = a^h, \end{aligned} \tag{3.5}$$

where a^h denotes the scaled matrix:

$$a^h = \{a_{ij}/(h_i h_j)\}.$$

Equality (3.5) means that a^h belongs to the cone generated by $\{\xi\xi^\mathsf{T}; \xi \in \mathcal{S}\}$,

$$\mathcal{C}(\mathcal{S}) = \left\{ \sum_{\xi \in \mathcal{S}} \gamma_\xi \xi\xi^\mathsf{T}, \gamma \in \mathbb{R}_+^{|\mathcal{S}|} \right\}.$$

A natural choice for $\mathcal{S}$ is the following:

$$\mathcal{S} = \mathcal{S}_p = \{(\xi_1, \xi_2) \in \mathbb{Z} \times \mathbb{N}; \max(|\xi_1|; \xi_2) \leq p; (|\xi_1|, \xi_2) \text{ irreducible}\},$$

for $p \geq 1$, and the correspondent cones $\mathcal{C}(\mathcal{S}_p)$. These cones have the following property:

$$\mathcal{C}(\mathcal{S}_1) \subset \mathcal{C}(\mathcal{S}_2) \subset \cdots \subset \mathcal{C}(\mathcal{S}_p) \subset \cdots \subset \mathcal{M}_+^\#,$$

where $\mathcal{M}_+^\#$ denotes the set of symmetric positive matrices. We can represent these matrices

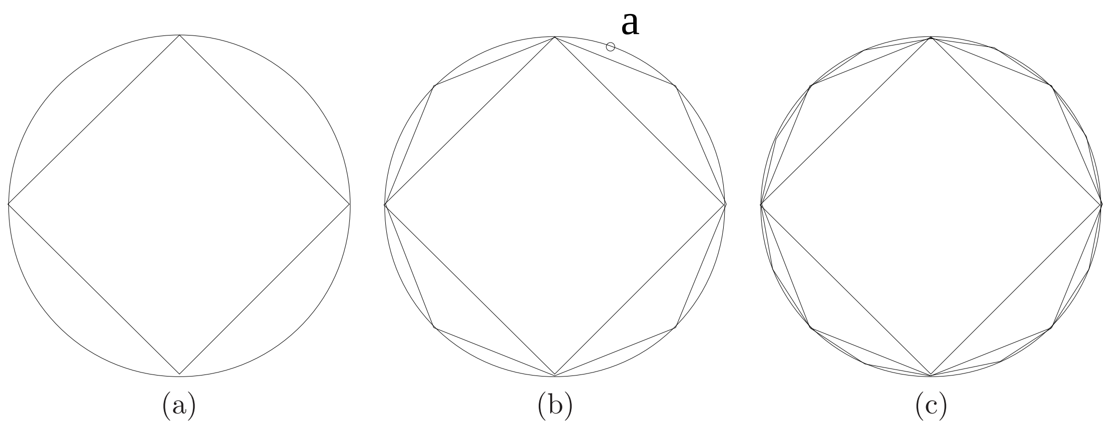


Figure 2.1: (a) Symmetric semi-definite positive matrix with trace equal to 1 and cone of diagonal dominant matrix. (b) Cone $\mathcal{C}(\mathcal{S}_1)$, a is on the border of the semi-definite positive matrix. (c) Cone $\mathcal{C}(\mathcal{S}_2)$.

in $\mathbb{R}^3$ using the following coordinates:

$$z_1 = a_{11} \ , \ z_2 = \sqrt{2} a_{12} \ , \ z_3 = a_{22}.$$

The cone of symmetric matrices is represented in figure 3 (a), together with the cone $\mathcal{C}(\mathcal{S}_1)$ of diagonally dominant matrices. One can make a cut of these cones, considering matrices of a given constant trace. Then, we will use the following set of coordinates:

$$w_1 = \frac{z_1 - z_3}{\sqrt{2}} \ , \ w_2 = z_2 \ , \ w_3 = \frac{z_1 + z_3}{\sqrt{2}}.$$

The cut will be made for constant w_3, the axis of the two dimensional representation being w_1 and w_2. The cuts of cones $\mathcal{C}(\mathcal{S}_1)$, $\mathcal{C}(\mathcal{S}_2)$, and $\mathcal{C}(\mathcal{S}_3)$ with the plan of trace 1 matrices are represented on figure 3. Unfortunately, even for a big order $p >> 1$, the matrix a^h does

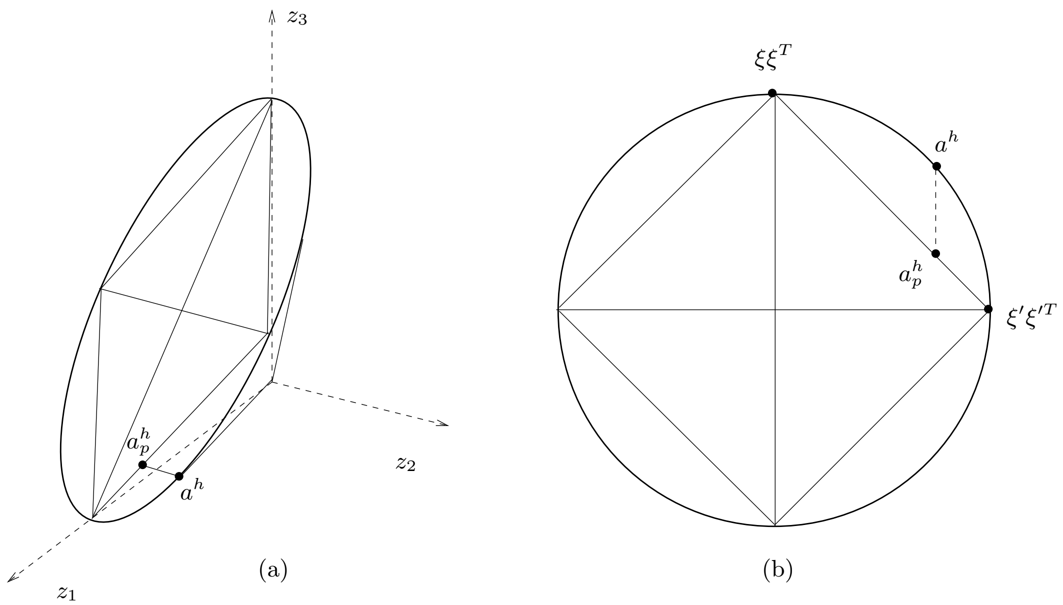


Figure 2.2: (a) Cone of positive definite matrices, embedding the cone of diagonally dominant matrices $\mathcal{C}(\mathcal{S}_1)$, and projection of a matrix a^h on $\mathcal{C}(\mathcal{S}_1)$. (b) Same figure, where we draw the cut of the cone for matrices of trace $\mathrm{Tr}(a_h)$.

not satisfy necessarily the strong consistency (3.5).

Moreover a^h is a rank one matrix and it is degenerated. This fact implies two possibilities:

• The direction of diffusion $\begin{pmatrix} \alpha_1 \sigma x \\ \alpha_2 \eta \end{pmatrix}$ points toward a point of the grid. This situation happens if the slope is a rational number r/q (with $r \in \mathbb{Z}$ and $q \in \mathbb{N}^*$). Then we consider the vector $\xi_{r,q} = (r\ q)^\mathsf{T}$, and we can write

$$a^h = \gamma_{\xi_{r,q}}^{\alpha_1,\alpha_2} \xi_{r,q} \xi_{r,q}^\mathsf{T}.$$

• The second possibility is that the direction of the diffusion $\begin{pmatrix} \alpha_1 \sigma x \\ \alpha_2 \eta \end{pmatrix}$ has a real slope. In this case, we approximate a^h by its projection a_p^h parallely to the z_2 axis (see figure 3) into one of the cones $\mathcal{C}(\mathcal{S}_p)$, the order p being the order of neighbouring points allowed to enter in the scheme (of course, this order depends on where we are situated on the grid and on the direction of the diffusion). Note that the quantity $a_{11} - a_{22}$ is conserved and as the trace does not change, and we obtain that a_{11} and a_{22} are invariant by this projection. Only a_{12} is modified.

Remark 3.1. *As we can see in Figure 3 (b), matrix a^h belongs to the border of the cone $\mathcal{M}_+^\#$ (the cone of symmetric semi-definite positive matrices), and then there exist two vectors $\xi_{p',q'}$ and $\xi_{p",q"}$ on $\mathcal{S}_p$, such that we can project a^h on the hyperplane generated by $\xi_{p',q'}\xi_{p',q'}^\mathsf{T}$ and $\xi_{p",q"}\xi_{p",q"}^\mathsf{T}$. Then, we can write the projection of a^h as follows:*

$$a_p^h = \gamma_{\xi_{p',q'}}^{\alpha_1,\alpha_2} \xi_{p',q'} \xi_{p',q'}^\mathsf{T} + \gamma_{\xi_{p",q"}}^{\alpha_1,\alpha_2} \xi_{p",q"} \xi_{p",q"}^\mathsf{T}, \tag{3.6}$$

where $\gamma_\xi^{\alpha_1,\alpha_2}$ *are positive coefficients, and moreover*

$$\gamma_{\xi_{p',q'}}^{\alpha_1,\alpha_2} + \gamma_{\xi_{p",q"}}^{\alpha_1,\alpha_2} \leq \mathrm{tr}(a_p^h).$$

As studied in [17], the generation of the directions $\xi_{p',q'}$ *and* $\xi_{p",q"}$*, can be performed (in effective way) in* $O(p)$ *operations, by using Stern-Brocot algorithm [41].*

Remark 3.2. *This is not the same projection as in [17], where an orthogonal projection used. This modification is important to obtain the global convergence of our scheme (see section 5).*

Remark 3.3. *The choice of the order* p *depends on where we are situated on the grid. For instance, if we consider a point* (x, y) *in the middle of the grid, and we want to discretize* $a \cdot D^2\phi(t, x, y)$*, we can follow the direction of diffusion and choose the biggest order of discretization* p*, because more* p *is bigger and better is the approximation of the scaled covariance matrix* a^h*. On the other hand, if we consider a point* (x, y) *near to the boundary, it can often happen that following the direction of the diffusion, we involve in the discretization some points which are out of the grid. In this case the choice of* p *is not free, and we refer to the Appendix for a detailed discussion of this case.*

Remark 3.4. *In all the decompositions, the coefficients* $\gamma_\xi^{\alpha_1,\alpha_2}$ *and also the vectors* ξ *are in terms of* (t, x, y)*. Sometimes, for simplicity of notations we do no specify this dependence.*

Error projection for scaled covariance matrices. For a symmetric matrix b of dimension 2 we use the Frobenius norm $\|b\| = (\sum_{i,j=1,2} b_{ij}^2)^{1/2}$. Let $p_{\max}$ the maximum order that we can consider for the discretization, and let us consider the projection b' of a general matrix $b \in \mathcal{M}_+^\#$ on the cone $\mathcal{C}(\mathcal{S}_{p_{\max}})$. It is easy to prove by geometrical considerations that the projection b' of a general matrix $b \in \mathcal{M}_2$, on a hyperplane of $\mathcal{C}(\mathcal{S}_{p_{\max}})$ spanned by $\xi\xi^T$ and $\xi'(\xi')^T$ is such that

$$\|b - b'\| \leq \widehat{\xi, \xi'}\|b\|, \tag{3.7}$$

where $\widehat{\xi, \xi'}$ denotes the length of the arc (ξ, ξ'), see figure 3 (b). From this inequality and [17, Lemma 6.1], the projection error will be bounded by $\epsilon_p||b||$, where

$$\epsilon_p = \frac{\pi}{2p_{\max}}.$$

Therefore, the error projection is guaranteed to be at most equal to ε (for any $\varepsilon > 0$), if we choose $p_{\max} \geq p_\varepsilon$, where

$$p_\varepsilon := \frac{\pi}{2\varepsilon}. \tag{3.8}$$

3.1 The discrete equation

From now on, by $\lceil r \rceil$ we denote the smallest integer greater than r, and we fix $h_1 = h_2 = h$, the space step size[5], $p_{\max} \in \mathbb{N}$ the maximal order of grid points allowed to enter in the

[5] We set $h_1 = h_2 = h$ to simplify the analysis.

scheme, and Δt the time step sizes. Set $\rho = (p_{\max}, h, \Delta t)$, and define the scheme S^ρ (given in a general setting) as follows. Let $\phi : [0,T] \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ and

$$S^\rho(t,x,y,r,\phi) \;:=\; \min_{\alpha_1^2+\alpha_2^2=1} \Big\{ -\alpha_1^2 \frac{\phi(t+\Delta t,x,y)-r}{\Delta t} + \alpha_1^2 \mu \frac{r-\phi(t,x,y-h)}{h}$$
$$-\frac{1}{2} \sum_{\xi \in \mathcal{S}(x,y)} \gamma_\xi^{\alpha_1,\alpha_2}(t,x,y) \big[\phi(t,x-\xi_1 h, y-\xi_2 h) - 2r + \phi(t,x+\xi_1 h, y+\xi_2 h)\big] \Big\}, \tag{3.9a}$$

for $(t,x,y) \in [0,T) \times (0,\infty)^2$, where

$$\begin{array}{l} \mathcal{S}(x,y) := \mathcal{S}_p \text{ with } p = \min(p_{max}, \lceil x/h \rceil, \lceil y/h \rceil), \\ \sum_{\xi \in \mathcal{S}(x,y)} \gamma_\xi^{\alpha_1,\alpha_2}(t,x,y) \xi \xi^\mathsf{T} = a_p^h(t,x,y), \end{array} \tag{3.9b}$$

the projection of the scaled covariance matrix a^h on $\mathcal{C}(\mathcal{S}_p)$ ($a^h = a/h^2$). In particular, $p = p_{max}$ if $x - p_{max}h \geq 0$ and $y - p_{max}h \geq 0$ (points in the interior of the domain), otherwise $p = \min(\lceil x/h \rceil, \lceil y/h \rceil)$ (points near to the boundary).

Now the discrete scheme for (3.1) is:

$$S^\rho(t,x,y,v_h(t,x,y),v_h) = 0, \tag{3.10a}$$

for $(t,x,y) \in [0,T) \times (0,\infty)^2$, and with the boundary conditions:

$$v_h(T,x,y) = g(x),\ \forall\ (x,y) \in [0,\infty)^2, \tag{3.10b}$$
$$v_h(t,x,0) = g(x),\ \forall\ (t,x) \in [0,T] \times [0,\infty) \tag{3.10c}$$
$$v_h(t,0,y) = g(0),\ \forall\ (t,y) \in [0,T] \times [0,\infty) \tag{3.10d}$$

(the solution v_h will stand for an approximation of the value function ϑ.)

Remark 3.5. *It is clear that if $p_{\max}$ is not linked to the step size h, then* (3.10) *is a discrete scheme for the HJB equation with the covariance matrix $a_p = h^2 a_p^h$ instead of a:*

$$\min_{\alpha_1^2+\alpha_2^2} \Big\{ -\alpha_1^2 \frac{\partial \phi}{\partial t}(t,x,y) + \alpha_1^2 \mu \frac{\partial \phi}{\partial y}(t,x,y) - \frac{1}{2} \mathrm{tr}[a_p D^2 \phi(t,x,y)] \Big\} = 0. \tag{3.11}$$

In what follows (subsection 3.2, and section 4), we will prove that the scheme (3.10) satisfies the following properties:

(S1) **Monotonicity:** $S^\rho(t,x,y,r,u) \geq S^\rho(t,x,y,r,v)$,
for all $r \in \mathbb{R}$, $x,y \in \mathbb{R}_+^*$, $u,v \in C([0,T] \times [0,\infty)^2)$ such that $u \leq v$ in $[0,T] \times [0,\infty)^2$.

(S2) **Stability:** For all $\rho = (h, \Delta t, p_{\max}) \in (\mathbb{R}_+^*) \times (0,T) \times \mathbb{N}^*$, there exists a bounded solution v_h of (3.10).

(S3) **Consistency:** There exists a constant $C_1 > 0$, and a constant such that, for every $\phi \in C^n([0,T] \times [0,\infty)^2)$, $n \geq 4$, with bounded derivatives,

$$\begin{aligned} &\Big| \min_{\alpha_1^2+\alpha_2^2} \Big\{ -\alpha_1^2 \frac{\partial \phi}{\partial t}(t,x,y) + \alpha_1^2 \mu \frac{\partial \phi}{\partial y}(t,x,y) - \frac{1}{2} \mathrm{tr}[a \cdot D^2 \phi(t,x,y)] \Big\} \\ &- S^\rho(t,x,y,\phi(t,x,y),\phi) \Big| \\ &\leq C_1(|\partial_t^2 \phi|_0 \Delta t + \mu |D_y^2 \phi|_0 h) + 16\sqrt{2} p_{max}^2 \|a\| |D^4 \phi|_0 h^2 + \varepsilon_p(t,x,y) |D^2 \phi|_0, \end{aligned} \tag{3.12}$$

where a_p is the projection of a on $\mathcal{C}(\mathcal{S}_p)$, for $p = \min(p_{max}, \lceil x/h \rceil, \lceil y/h \rceil)$, and $\varepsilon_p(t,x,y)$ is the projection error such that $\varepsilon_p = ||a - a_p||$ if $p = p_{max}$, and $\varepsilon_p = CK(x,y)h$ otherwise, where C depends on the Lipschitz constant of σ^2.

3.2 The consistency property

We start by proving the consistency property (S3). Consider a function $\phi \in C^n([0,T] \times [0,+\infty)^2)$, with bounded derivatives and compute (first) the difference term:

$$\begin{aligned} &\Bigg| \min_{\alpha_1^2+\alpha_2^2} \Big\{ -\alpha_1^2 \frac{\partial \phi}{\partial t}(t,x,y) + \alpha_1^2 \mu \frac{\partial \phi}{\partial y}(t,x,y) - \frac{1}{2}\mathrm{tr}[a_p \cdot D^2\phi(t,x,y)] \Big\} \\ &-S^\rho(t,x,y,\phi(t,x,y),\phi) \Bigg|, \end{aligned} \tag{3.13}$$

for the HJB-equation with the matrix a_p instead of a. For the derivatives on t and on y we just apply a Taylor development to obtain the bound terms $|\partial_t^2\phi|_0 \Delta t$ and $\mu |D_y^2 \phi|_0 h$. Consider now the diffusion term: by a Taylor development, we get (for $\xi \in \mathcal{S}$):

$$\begin{aligned} D^2\phi(h\xi, h\xi) - \Delta_\xi \phi &\leq 2h^4 \textstyle\sum_{k=0}^4 \xi_1^k \xi_2^{4-k} \frac{\partial^4 \phi}{\partial x^k \partial y^{4-k}}, \\ &\leq 4h^4 ||\xi||^4 |D^4\phi|_0, \end{aligned}$$

where $D^4\phi = \sum_{k=0}^4 \frac{\partial^4 \phi}{\partial x_1^k \partial x_2^{4-k}}$, and the last inequality follows from the fact that $\sum_{k=0}^4 \xi_1^k \xi_2^{4-k} \leq 2||\xi||^4$. Moreover, from (3.6), we can deduce that

$$0 \leq \gamma_\xi^{\alpha_1,\alpha_2} \leq \frac{\mathrm{tr}(a_p^h)}{\|\xi\|^2},$$

for every ξ which appear in the decomposition of a_p^h. Then, for the global diffusion term we obtain

$$\begin{aligned} \mathrm{tr}[a_p D^2\phi(t,x,y)] - \textstyle\sum_{\xi \in \mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \Delta_\xi \phi(t,x,y) &\leq 2\mathrm{tr}(a_p^h)|D^4\phi|_0 h^4 \textstyle\sum_{\xi\in\mathcal{S}} \|\xi\|^2 \\ &\leq 8\mathrm{tr}(a_p^h)|D^4\phi|_0 h^4 p_{max}^2 \\ &\leq 8\mathrm{tr}(a_p)|D^4\phi|_0 h^2 p_{max}^2 \end{aligned}$$

where the last inequality follows from the fact that $\xi_i \leq p_{max}$, for $i = 1,2$. We are now looking for a bound of $\mathrm{tr}(a_p)$ which depends on (t,x,y). It is easy to see that $\mathrm{tr}(a_p) \leq \sqrt{2}\|a_p\|$, and moreover, by 3.7, we can show that

$$\|a_p\| \leq 2\|a\|, \tag{3.14}$$

where $\|a\|$ depends on t,x,y.

Therefore, we obtain

$$\mathrm{tr}[a_p D^2\phi(t,x,y)] - \sum_{\xi \in \mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \Delta_\xi \phi(t,x,y) \leq 16\sqrt{2}\|a\| |D^4\phi|_0 h^2 p_{max}^2.$$

Then we can conclude that

$$(3.13) \quad \leq \quad C_1(|\partial_t^2\phi|_0 \Delta t + \mu |D_y^2\phi|_0 h) + 16\sqrt{2} p_{max}^2 \|a\| |D^4\phi|_0 h^2. \tag{3.15}$$

On the other hand,

- For the points (x, y) such that $x - p_{max}h \geq 0$, and $y - p_{max}h \geq 0$,

$$\left|\mathrm{tr}[a \cdot D^2\phi(t,x,y)] - \mathrm{tr}[a_p \cdot D^2\phi(t,x,y)]\right| \leq \|a - a_p\||D^2\phi|_0 \leq \varepsilon_p|D^2\phi|_0.$$

  Let us note that for a fixed ε_p depending on the space step h, following (3.8) we can give a condition on p_{max} to obtain an error which is at most equal to ε_p.

- For the points such that $x < p_{max}h$ or $y < p_{max}h$, using the equivalent formulation of remark 2.11 with $\eta(t, y) = \min(1, y)$, and the Lipschitz property of σ^2, one gets

$$||a|| \leq C(x^2y + y^2 + xy),$$

  where C depends on α_1,α_2 and on the Lipschitz constant of σ^2. Moreover, since we know that $p = \min(\lceil x/h\rceil, \lceil y/h\rceil)$, then taking $K(x, y) = (x \vee y + 1)^2$, we get $||a|| \leq CK(x, y)ph$. Then we obtain the following estimate:

$$\begin{aligned}\left|\mathrm{tr}[a \cdot D^2\phi(t,x,y)] - \mathrm{tr}[a_p \cdot D^2\phi(t,x,y)]\right| &\leq \|a - a_p\||D^2\phi|_0 \\ &\leq \epsilon_p^h||a|| \cdot |D^2\phi|_0 \\ &\leq 2CK(x,y)|D^2\phi|_0 h,\end{aligned}$$

  where the last inequality follows from the fact that $\epsilon_p \leq 2/(p)$.

This concludes the consistency property (S3), with $\varepsilon_p(t, x, y) = ||a - a_p||$ if $p = p_{max}$, and $\varepsilon_p(t, x, y) = CK(x, y)h^2$ if $p = \min(\lceil x/h\rceil, \lceil y/h\rceil)$.

We give the following result.

Proposition 3.17. *Suppose that*

$$p_{max} \sim \frac{C}{h^{\frac{2}{3}}},$$

then

$$\left|\min_{\alpha_1^2+\alpha_2^2}\left\{ -\alpha_1^2\frac{\partial\phi}{\partial t}(t,x,y) + \alpha_1^2\mu\frac{\partial\phi}{\partial y}(t,x,y) - \frac{1}{2}\mathrm{tr}[a \cdot D^2\phi(t,x,y)]\right\} \right.$$
$$\left. -S^\rho(t,x,y,\phi(t,x,y),\phi)\right| = O(h^{\frac{2}{3}}) + O(\Delta t).$$

for all $(t, x, y) \in [0, T] \times (\mathbb{R}_+^*)^2$.

Proof. The proof follows from the explicit form of the consistency property (S3). □

Remark 3.6. *In the case when the direction of diffusion points toward a point of the grid, the consistency remains the same, except for the error of projection which will be zero.*

4 Existence of the numerical solution

In this section we prove the well-posedness of the implicit scheme (3.10a) with boundary conditions (3.10b), (3.10c) and (3.10d), and show that it satisfies the required monotony and stability assumptions (S1)-(S2).

We recall that the grid is $G_h = \{(x_i, y_j), i, j \geq 0\} \subset \mathbb{R}_+^2$, where $x_i := ih_1$, $y_j := jh_2$.

We first start to initialize the scheme by

$$v_h(T,x,y) := g(x), \qquad (x,y) \in G_h.$$

Then, given $v_h(t+\Delta t,.)$ for some time t, we need to find $v_h(t,x,y)$ for $(x,y) \in G_h$ such that

$$\min_{\alpha_1^2+\alpha_2^2=1} \Big\{ \alpha_1^2 \frac{v_h(t,x,y) - v_h(t+\Delta t,x,y)}{\Delta t} + \alpha_1^2 \mu(t,y) \frac{v_h(t,x,y) - v_h(t,x,y-h_2)}{h_2}$$
$$- \frac{1}{2} \sum_{\xi \in \mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2} \Delta_\xi v_h(t,x,y) \Big\} = 0, \quad \forall (x,y) \in G_h,\ y>0, \tag{4.1}$$

and with the following "boundary conditions":

$$v_h(t,x,0) = g(x),\ \forall\, x \in h_1\mathbb{N} \tag{4.2}$$
$$v_h(t,.,.) \text{ bounded} \tag{4.3}$$

The scheme in abstract form. Since for all $(x,y) \in G_h$ with $y>0$, an optimal control (α_1,α_2) must be found, we introduce $S^1 := \{\alpha = (\alpha_1,\alpha_2), \alpha_1^2+\alpha_2^2 = 1\}$ and

$$\mathcal{A} := (S^1)^{\mathbb{N}\times\mathbb{N}^*}$$

the set of controls associated to the grid mesh $(x_i,y_j)_{i\geq 0,\, j\geq 1}$.

The scheme can then be expressed in the following abstract form: find $X := v_h(t,.,.) \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}^*}$, bounded, such that

$$\min_{w\in\mathcal{A}} \Big(A(w)X - b(w) \Big) = 0, \tag{4.4}$$

where $A(w)$ is a linear operator on $\mathbb{R}^{\mathbb{N}\times\mathbb{N}^*}$, and $b(w)$ is a vector of $\mathbb{R}^{\mathbb{N}\times\mathbb{N}^*}$, and are made precise below.

Definition of the matrix $A(w)$ and vector $b(w)$: Let $X = (X_{ij})_{i\geq 0,\, j\geq 1}$, (resp. $w = (\alpha_{ij})_{i\geq 0,\, j\geq 1}$, with $\alpha_{ij} = (\alpha_{ij,1},\alpha_{ij,2})$) be values (resp. controls) corresponding to the mesh points (x_i,y_j) of G_h. Then

- $A(w)$ is an infinite matrix determined by $\forall X$, $\forall i \geq 0$, $\forall j \geq 1$,

$$\begin{aligned}(A(w)X)_{ij} \;:=\; & \frac{\alpha_{ij,1}^2}{\Delta t} X_{ij} + \alpha_{ij,1}^2 \mu(t,y_j) \frac{1}{h_2}(X_{ij} - (1-\kappa_{j-1})X_{i,j-1}) \\ & + \frac{1}{2} \sum_{\xi=(\xi_1,\xi_2)\in\mathcal{S}} \gamma_\xi^{\alpha_{ij}} (-(1-\kappa_{j-\xi_2})X_{i-\xi_1,j-\xi_2} + 2X_{ij} - X_{i+\xi_1,j+\xi_2})\end{aligned}$$

  where $\kappa_k := 1$ if $k=0$ and $\kappa_k := 0$ if $k \neq 0$.

- $b(w)$ is defined by

$$\begin{aligned}b_{i,j}(w) \;:=\; & \frac{\alpha_{ij,1}^2}{\Delta t} v_h(t+\Delta t,x_i,y_j) + \alpha_{ij,1}^2 \frac{\mu(t,y_j)}{h_2} \kappa_{j-1} g(x_i) \\ & + \frac{1}{2} \sum_{\xi=(\xi_1,\xi_2)\in\mathcal{S}} \gamma_\xi^{\alpha_{ij}} \kappa_{j-\xi_2} g(x_{i-\xi_1})\end{aligned} \tag{4.5}$$

  where $v_h(t+\Delta t,x,y)$ is the solution at the previous time step and is assumed bounded.

We shall also denote

$$\delta_{ij}(w) := \frac{\alpha_{ij,1}^2}{\Delta t} + \frac{\alpha_{ij,1}^2}{h_2}\mu(t,y_j)\kappa_{j-1} + \frac{1}{2}\sum_{\xi=(\xi_1,\xi_2)\in\mathcal{S}} \gamma_\xi^{\alpha_{ij}}\kappa_{j-\xi_2}.$$

Remark 4.1. *The matrix $A(w)$ is $\delta(w)$-diagonal dominant in the following sense:*

$$A_{(i,j),(i,j)}(w) = \delta_{ij}(w) + \sum_{(k,\ell)\neq(i,j)} |A_{(i,j),(k,\ell)}(w)|$$

Remark 4.2. *Note that in the case no border points $y=0$ are involved (i.e. when $j > p_{max}$), we have the more simple expressions:*

$$\begin{aligned}(A(w)X)_{ij} \quad := \quad & \frac{\alpha_{ij,1}^2}{\Delta t}X_{ij} + \frac{\alpha_{ij,1}^2}{h_2}\mu(t,y_j)(X_{ij}-X_{i,j-1}) \\ & +\frac{1}{2}\sum_{\xi=(\xi_1,\xi_2)\in\mathcal{S}} \gamma_\xi^{\alpha_{ij}}(-X_{i-\xi_1,j-\xi_2} + 2X_{ij} - X_{i+\xi_1,j+\xi_2}).\end{aligned}$$

and

$$b_{i,j}(w) := \frac{\alpha_{ij,1}^2}{\Delta t}v_h(t+\Delta t, x_i, y_j), \qquad \delta_{ij}(w) := \frac{\alpha_{ij,1}^2}{\Delta t}.$$

Remark 4.3. *Note that on the boundary $x=0$, if we assume that $v_h(t+\Delta t,0,y) = g(0)$ then the scheme reads*

$$\begin{aligned}\min_{\alpha_1^2+\alpha_2^2=1} \Big\{ & \alpha_1^2\frac{v_h(t,0,y)-g(0)}{\Delta t} + \alpha_1^2\mu(t,y)\,\frac{v_h(t,0,y)-v_h(t,0,y-h_2)}{h_2} \\ & +\frac{1}{2}\alpha_2^2(-v_h(t,0,y-h_2) + 2v_h(t,0,y) - v_h(t,0,y+h_2))\Big\} = 0, \quad \forall y\in h_2\mathbb{N}^* \quad (4.6)\end{aligned}$$

and with $v_h(t,0,0) = g(0)$. One can show that $v_h(t,0,y) = const = g(0)$ is the only bounded solution of (4.6) *(using the results of Lemma 8.14, Prop. 8.20 and Prop. 9.22). Hence by recursion we see that $v_h(t,0,y) = g(0)$ for all t and $y\in h_2\mathbb{N}$. In order to simplify the presentation of $A(w)$ and $b(w)$ we have preferred not to add this knowledge in a boundary condition at $x=0$.*

Preliminary results. In order to find a solution of (4.4), we first consider the linear system

$$A(w)X = b(w),$$

for a given $w\in\mathcal{A}$. For clarity, some specific results for such systems have been postponed to an appendix. We can check that $(A(w),b(w))$ satisfy all the assumptions of Proposition 8.21. In particular, we obtain that $A(w)$ is a *monotone matrix*, in the sense that if $X=(X_{i,j})_{i\geq 0,j\geq 1}$ is bounded (or bounded from below) and such that

$$\forall i\geq 0,\ \forall j\geq 1, \quad \delta_{ij}(w) = 0\ \Rightarrow (A(w)X)_{ij} = 0, \tag{4.7}$$

then

$$A(w)X\geq 0 \Rightarrow\ X\geq 0.$$

Here (4.7) is equivalent to

$$\forall i \geq 0,\ j \geq 1, \quad \alpha_{ij,1} = 0 \ \Rightarrow -X_{i,j-1} + 2X_{ij} - X_{i,j+1} = 0.$$

Since $b(w)$ satisfies $\delta_{ij}(w) = 0 \Rightarrow b_{i,j}(w) = 0$, and that

$$\max_{i,j;\delta_{ij}(w)>0} \frac{|b_{ij}(w)|}{\delta_{ij}(w)} \leq \max(||v_h(t+\Delta t,.,.)||_\infty,\ ||g||_\infty),$$

we also obtain by Proposition 8.21 (ii) that there exists a unique bounded X such that $A(w)X = b(w)$, and satisfying furthermore

$$||X||_\infty := \max_{i\geq 0, j\geq 1} |X_{ij}| \leq \max(||v_h(t+\Delta t,.,.)||_\infty,\ ||g||_\infty).$$

Howard algorithm We can now consider the following Howard algorithm for solving (4.4).

Let $w^0 \in \mathcal{A}$ be a given initial control value

Iterate for $k \geq 0$

- Find X^k bounded, such that $A(w^k)X^k = b(w^k)$.
- $w^{k+1} := \mathrm{argmin}_{w\in\mathcal{A}}(A(w)X^k - b(w))$.

In the second step note that the minimization is done component by component, since $(A(w)X^k - b(w))_{ij}$ depends only of the control α_{ij}; the minimum is also well defined since the control set S^1 for α_{ij} is compact.

Then we have the following result, whose proof is postponed to the appendix.

Proposition 4.18. *There exists a unique bounded solution X to the problem*

$$\min_{w\in\mathcal{A}}(A(w)X - b(w)) = 0,$$

and the sequence X^k converges pointwisely towards X, i.e., $\lim_{k\to\infty} X^k_{ij} = X_{ij}$ $\forall i,j \geq 0$.

Stability and monotonicity. First, the convergence thus leads also to the bound $||v_h(t,.)||_\infty = ||X||_\infty \leq \max(||v_h(t+\Delta t,.)||_\infty,\ ||g||_\infty)$. Hence by recursion we obtain $||v_h(t,.)||_\infty \leq ||g||_\infty$, which shows the stability of the scheme.

Then, the monotonicity is also obtained directly from the definition of the scheme.

Remark 4.4. *Note that we have the following stronger monotonicity result: if $v_h^1(t+\Delta t)$ and $v_h^1(t+\Delta t)$ are two bounded vectors defined on the grid, and X^1 and X^2 denotes the two corresponding solutions of* (4.4)*, then*

$$v_h^1(t+\Delta t,.) \leq v_h^2(t+\Delta t,.) \quad \Rightarrow \quad X^1 \leq X^2.$$

To see this, let us denote $b^q(w)$, for $q = 1, 2$, the vectors corresponding to $v_h^q(t+\Delta t)$ as defined in (4.6)*. We note that $b^1(w) \leq b^2(w)$, $\forall w \in \mathcal{A}$. Let w^1 be an optimal control for X^1. Then*

$$\begin{aligned} A(w^1)X^1 - b^1(w^1) &= 0 = \min_{w\in\mathcal{A}}(A(w)X^2 - b^2(w)) \\ &\leq A(w^1)X^2 - b^2(w^1) \\ &\leq A(w^1)X^2 - b^1(w^1), \end{aligned}$$

and thus $A(w^1)(X^2 - X^1) \geq 0$. By the monotonicity property of $A(w^1)$ and the fact that if $\delta_{ij}(w^1) = 0$ then $b^2_{ij}(w^1) - b^1_{ij}(w^1) = 0$, we conclude to $X^1 \leq X^2$.

Remark 4.5. *Note that the stability and monotonicity results are obtained inconditionnaly with respect to the mesh sizes h_1, h_2 and $\Delta t > 0$.*

5 Convergence

Since the scheme is monotone, stable and consistent, we can use the same arguments as in [10, Theorem 2.1] to conclude the convergence of ϑ_h toward ϑ, taking into account the comparison principle Theorem 2.5.

In order to prove the convergence, we first note that the following type of discrete comparison principle holds for the scheme.

Lemma 5.13. *Let $Y = Y_{h,\Delta t}(t,x,y)$ be defined on $(x,y) \in G_h$ and for $T - t \in \Delta t\mathbb{N}$. Suppose that Y is a super-solution of the scheme (resp sub-solution of the scheme), in the following sense:*
(i) $\forall t+\Delta t \leq T$, $(x,y) \in G_h$, $y > 0$, $S^\rho(t,x,y,Y(t,x,y),Y) \geq 0$ (resp. $S^\rho(t,x,y,Y(t,x,y)) \leq 0$),
(ii) $\forall (x,y) \in G_h$, $Y(T,x,y) \geq g(x)$ (resp $Y(T,x,y) \leq g(x)$),
(iii) $\forall t \leq T$, $(x,y) \in G_h$, $Y(t,x,0) \geq g(x)$ (resp $Y(t,x,0) \leq g(x)$),
(iv) $Y(t,x,y)$ is bounded from below (resp. from above).
Then $Y \geq v_h$ (resp $Y \leq v_h$), where $v_h = v_h(t,x,y)$ are the scheme values.

Proof. Indeed the proof can be obtained by recursion (using $Y(t+\Delta t,.) \geq v_h(t+\Delta t,.)$ to show that $Y(t,.) \geq v_h(t,.)$) following the same arguments as in Remark 4.4. In order to conclude from $A(w_1)(Y(t,.) - v_h(t,.)) \geq 0$ to $Y(t,.) - v_h(t,.) \geq 0$ (for a given control w_1), we use the fact that $Y(t,.) - v_h(t,.)$ is bounded from below and Prop 8.21 1). The proof for the sub-solution is similar. □

We can give now the convergence result.

Theorem 5.1. *We assume (A1)-(A3) and that g is C^2-regular and such that $-x^2 g''(x)$ be bounded from below. Suppose also that $p_{max} \to +\infty$ and $p_{max} = o(\frac{1}{h}$. Then the scheme converges locally uniformly to ϑ when $h, \Delta t \to 0$.*

Proof. In the following when we denote $h \to 0$ we also mean that $\Delta t \to 0$. Let $\bar{v}$ and $\underline{v}$ be defined by

$$\overline{v}(t,x,y) = \limsup_{h,\Delta t \to 0,(t',x',y')\to(t,x,y)} v_h(t',x',y'),$$
$$\underline{v}(t,x,y) = \liminf_{h,\Delta t \to 0,(t',x',y')\to(t,x,y)} v_h(t',x',y'),$$

(The function $v_h(t,x,y)$ defined for (x,y) in the grid G_h and for $T - t = n\Delta t$ can be extended to $[T,0] \times \mathbb{R}^+ \times \mathbb{R}^+$ by a P0 interpolation.) As in [10, Theorem 2.1], using properties (S1-S3) of the scheme, we can prove that $\bar{v}$ and $\underline{v}$ are respectively bounded viscosity sub- and super-solution of (3.1). If the following inequalities hold:

$$\overline{v}(T,x,y) \leq g(x) \leq \underline{v}(T,x,y) \tag{5.1}$$
$$\overline{v}(t,x,0) \leq g(x) \leq \underline{v}(t,x,0) \tag{5.2}$$

then, by the comparison principle (Theorem 2.5) we obtain $\overline{v} \leq \underline{v}$, hence $\overline{v} = \underline{v}$ and the convergence of v_h towards the unique viscosity solution of (3.1), i.e. ϑ.

Step 1: $\underline{v}(T,x,y) \geq g(x)$, and $\underline{v}(t,x,0) \geq g(x)$.
Considering $Y(t,x,y) = g(x)$, we see that Y is a sub-solution of the scheme (in the sense of Lemma 5.13). Hence $v_h \geq Y$ and we deduce the two inequalities $\underline{v}(T,x,y) \geq g(x)$ and $\underline{v}(t,x,0) \geq g(x)$.

Step 2: $\bar{v}(T,x,y) \leq g(x)$, and $\bar{v}(t,x,0) \leq g(x)$.
Let $A,B,C \geq 0$ and $L \geq 0$ be some constants such that $g''(x) \leq A$ for all $x \in [0,1]$, $B \geq \frac{1}{2}x^2 g''(x)$, $C > (x+1)g''(x)$ and $\sigma^2(t,y) \leq Ly$ for all positive x,y. Let K be a constant such that $K \geq (A \vee B + C)L$, and, for $t \in [0,T]$, let

$$Y(t,x,y) := K\,(T-t)\,y + g(x).$$

For some point $(t,x,y) \in \Delta t\mathbb{N} \times G_h$, fix the value of the controls α_1, α_2. We get that the value of the discrete operator can be decomposed as:

$$\begin{aligned} S(t,x,y,Y(t,x,y),Y) = \qquad & S(t,x,y,Y(t,x,y),Y) - \Big(-\alpha_1^2 \frac{\partial Y}{\partial t} - \alpha_2^2 \mu \frac{\partial Y}{\partial x} \\ & -\frac{1}{2} tr(a_p D^2 Y(t,x,y)) \Big) + \Big(\frac{1}{2} tr(a D^2 Y(t,x,y)) \\ & -\frac{1}{2} tr(a_p D^2 Y(t,x,y)) \Big) + H(t,x,y,Y(t,x,y),Y) \end{aligned}$$

where the diffusion matrix a and its projection a_p can be written

$$a = \begin{pmatrix} \sigma^2(t,y)x^2 & \xi\sigma(t,y)x \\ \xi\sigma(t,y)x & \xi^2 \end{pmatrix} \quad , \quad a_p = \begin{pmatrix} \sigma^2(t,y)x^2 & b \\ b & \xi^2 \end{pmatrix}$$

for some b. Consequently, we get that the projection error is null since:

$$\frac{1}{2} tr(a D^2 Y(t,x,y)) - \frac{1}{2} tr(a_p D^2 Y(t,x,y)) = (\xi\sigma(t,y)x - b)\,\frac{\partial^2 Y}{\partial x \partial y} = 0,$$

as $\frac{\partial^2 Y}{\partial x \partial y} = 0$. Moreover, let the discretization error Ψ be defined by:

$$\begin{aligned} \Psi \;=\; & S(t,x,y,Y(t,x,y),Y) \\ & - \left(-\alpha_1^2 \frac{\partial Y}{\partial t} - \alpha_2^2 \mu \frac{\partial Y}{\partial y} - \frac{1}{2} tr(a_p D^2 Y(t,x,y)) \right). \end{aligned}$$

Since $\partial_t^2 Y = D_y^2 Y = 0$, and using the fact that the only non-zero term in the matrix $D^4 Y$ is $\frac{\partial^4 Y}{\partial x^4}$, we get that the error is only due to the discretization of $\frac{\partial^2 Y}{\partial x^2}$. Therefore we get as g is $\mathcal{C}^2([0,+\infty))$:

$$\Psi(t,x,y,Y(t,x,y),Y) = \alpha_1^2 \frac{1}{2} x^2 \sigma^2(t,y)(g''(x) - g''(\theta_{x,h})),$$

for some $\theta_{x,h} \in [x - p_{max}h, x + p_{max}h]$. Hence, we get:

$$S(t,x,y,Y(t,x,y),Y) = H(t,x,y,Y(t,x,y),Y) + \alpha_1^2 \frac{1}{2} x^2 \sigma^2(t,y)(g''(\theta_{x,h}) - g''(x)).$$

On the other hand, by direct calculations we have:

$$H(t,x,y,Y(t,x,y),Y) = \alpha_1^2\left(Ky - \frac{1}{2}x^2\sigma^2(t,y)g''(x) + \mu(t,x)(T-t)\right).$$

Using that μ is a positive function we get:

$$\begin{aligned} S(t,x,y,Y(t,x,y),Y) &\geq \alpha_1^2\left(Ky - \frac{1}{2}x^2\sigma^2(t,y)g''(\theta_{x,h})\right) \\ &\geq \alpha_1^2 L\left(Cy - \frac{1}{2}(x^2-\theta_{x,h}^2)yg''(\theta_{x,h})\right) \\ &\geq \alpha_1^2 L\left(Cy - \frac{1}{2}(x-\theta_{x,h})(x+\theta_{x,h})yg''(\theta_{x,h})\right) \\ &\geq \alpha_1^2 yL\left(A - \frac{1}{2}p_{max}h|x+p_{max}h|g''(\theta_{x,h})\right) \geq 0 \end{aligned}$$

for all (t,x,y) as $p_{max}h < 1$ for sufficiently small h. Hence, Y satisfies the assumptions (i)-(iv) of Lemma 5.13, and thus $Y \geq v_h$. In particular,

$$\overline{v}(T,x,y) = \limsup_{h\to 0,(t',x',y')\to(T,x,y)} v_h(t',x',y') \leq \limsup_{h\to 0,(t',x',y')\to(T,x,y)} Y(t,x,y) = g(x).$$

We obtain $\overline{v}(t,x,0) \leq g(x)$ in the same way. □

6 Numerical results

The approximation scheme of Section 3, hereafter refered as the Implicit Euler scheme, or (IE) scheme, is now tested on some numerical examples. In all cases we have taken

$$\mu(t,y) = 0,$$

i.e. no transport term, because this is not the main difficulty of the equation. Also we fixed

$$\sigma(t,y) = \sqrt{y}. \tag{6.1}$$

All tests where done in Scilab (equivalent of Matlab), on a Pentium 4, 3Ghz computer.

6.1 Consistency test

Here we perform a verification of the consistency error of the spatial discretization. We consider the function

$$v(t,x,y) := 1 - e^{-x^2-y^2} + (T-t)^2, \tag{6.2}$$

and define f such that

$$f(t,x,y) := \tag{6.3}$$
$$\inf_{\alpha_1^2+\alpha_2^2=1}\left\{\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}^T\begin{pmatrix} -\frac{\partial v}{\partial t} + \mu(t,y)\frac{\partial v}{\partial y} - \frac{1}{2}\sigma^2(t,y)x^2\frac{\partial^2 v}{\partial x^2} & -\frac{1}{2}\sigma(t,y)xy\frac{\partial^2 v}{\partial x\partial y} \\ -\frac{1}{2}\sigma(t,y)xy\frac{\partial^2 v}{\partial x\partial y} & -\frac{1}{2}y^2\frac{\partial^2 v}{\partial y^2}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}\right\}.$$

Here f corresponds to the second member of (2.10) with $\vartheta = v$. An exact computation gives

$$
\begin{aligned}
f(t,x,y) &= \frac{1}{2}\Bigg[-\frac{\partial v}{\partial t} - \frac{1}{2}\sigma^2(t,x)x^2\frac{\partial^2 v}{\partial x^2} - \frac{1}{2}y^2\frac{\partial^2 v}{\partial y^2} \\
&\quad + \sqrt{\left(\frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2(t,x)x^2\frac{\partial^2 v}{\partial x^2} - \frac{1}{2}y^2\frac{\partial^2 v}{\partial y^2}\right)^2 + \left(\sigma(t,x)xy\frac{\partial^2 v}{\partial x\partial y}\right)^2}\Bigg].
\end{aligned} \tag{6.4}
$$

Following the definition of $\mathcal{S}^\rho$ in (1.6), and together with the fact that $\mu = 0$, we define here $\mathcal{S}^{\rho,N_u}$ such that:

$$
\begin{aligned}
\mathcal{S}^{\rho,N_u}(t,x,y,v(t,x,y),v) &= \min_{k=1,\ldots,N_u}\Bigg\{\alpha_{1,k}^2\frac{v(t,x,y)-v(t+\Delta t,x,y)}{\Delta t} \\
&\quad -\frac{1}{2}\sum_{\xi\in\mathcal{S}}\gamma_\xi^{\alpha_k}\Delta_\xi v(t,x,y)\Bigg\},
\end{aligned}
$$

where $\alpha_k = (\alpha_{1,k},\alpha_{2,k}) := e^{2i\pi k/(2N_u)}$ (note that it is sufficient to take half of the unit circle for the controls α in the definition of $\mathcal{S}^\rho$, and we do the same for $\mathcal{S}^{\rho,N_u}$).

Then we compute at time $t = T$, the value of

$$
\mathcal{S}^{\rho,N_u}(t,x,y,v(t,x,y),v) - f(t,x,y). \tag{6.5}
$$

The results are shown in Table 2.1, in L^∞ and L^2 norms. The space domain is $[0,x_{\max}]\times[0,y_{\max}]$ with $x_{\max} = y_{\max} = 3$, and we have used here Neumann boundary conditions on $x = x_{\max}$ and on $y = y_{\max}$.

number of space steps	N_u	p_{max}	L^2 error	L^∞ error	CPU time (seconds)
20× 20	20	2	0.0215	0.0396	0.23
40× 40	40	3	0.0094	0.0212	1.72
80× 80	80	4	0.0046	0.0121	14.98
160× 160	160	6	0.0020	0.0058	156.09

Table 2.1: Consistency error

Remark 6.1. *Contrary to the the definition of the projection of the matrix a in section 3, we chose an orthogonal projection of a on $\mathcal{C}(\mathcal{S}_p)$, as in [17]. Even if convergence is not proved in that case, it gives better numercal results. Note that, from [17], the projection error $\|a - a^p\|$ becomes bounded by $\frac{C}{p_{max}^2}$ for some constant C.*

Remark 6.2. *A key parameter for the discretization scheme is the maximum order $p_{\max}$ that we consider. From the expression of the theoretical consistency error, we take $p_{max} \simeq \frac{C}{\sqrt{h}}$ where C is a given constant, and h is the space step (which is the same here for the x or the y variable). We see that a small $p_{\max}$, as in Table 2.1, is numerically sufficient to obtain the desired consistency order (here $O(h)$).*

We obtain a consistency error that converges to zero with rate h in both L^∞ and L^2 norms. To this end we also found numerically that it was sufficient to increase the number of controls as the number of space steps (as is done in Table 2.1).

In table 2.2, we also see that the consistency error behaves as $O(\frac{1}{N_u})$ when N_u is sufficiently small. For large N_u, and fixed space steps, the error does not more diminish, because the spatial error dominates.

N_u	L^2 error	L^∞ error
5	0.035	0.051
10	0.014	0.024
20	0.010	0.022
40	0.009	0.021
80	0.009	0.021

Table 2.2: Error with varying number of controls N_u. Space steps 40×40 here.

6.2 Convergence test

Now we consider the time-dependant equation (2.11), with unkown ϑ and with a second member f defined by (6.4) and (6.2), and with terminal data $\vartheta(T,.,.) = v(T,.,.)$. In this case we know that the value of the solution is $\vartheta = v$.

The results are given in Table 2.3, where we test the Implicit Euler scheme and also the Crank-Nicolson scheme (second order in time, see [30]). We have used $T = 1$ with different time steps. We find that the (IE) scheme converges with rate $O(h) + O(\Delta t)$. The Crank-Nicolson (CN) scheme (see Remark 6.3 below) gives better numerical results with a similar computational cost, even if we have not proved its convergence.

In Table 2.4 (varying time steps), we see that the convergence rate is of order $O(\Delta t)^2$ as expected.

number of space steps	N_u	N	$p_{\max}$	L^2 error (EI)	L^∞ error (EI)	CPU time (seconds)	L^2 error (CN)	L^∞ error (CN)
20×20	20	20	2	0.0590	0.0822	98	0.0136	0.0333
40×40	40	40	3	0.0284	0.0367	946	0.0051	0.0117
80×80	80	80	4	0.0138	0.0178	10120	0.0023	0.0053

Table 2.3: Error for the Implicit Euler scheme and the Crank-Nicholson Scheme.

Remark 6.3. *The Crank-Nicolson scheme is defined here by the following implicit scheme:*

$$0 = \min_{k=1,\dots,N_u} \Bigg\{ \alpha_{1,k}^2 \frac{v(t,x,y) - v(t+\Delta t,x,y)}{\Delta t} + \frac{1}{2}\left(-\frac{1}{2}\sum_{\xi\in\mathcal{S}} \gamma_\xi^{\alpha_k} \Delta_\xi v(t,x,y) - \frac{1}{2}\sum_{\xi\in\mathcal{S}} \gamma_\xi^{\alpha_k} \Delta_\xi v(t+\Delta,x,y) \right) \Bigg\}$$

(note that in our test $\gamma_\xi^{\alpha_k}$ does not depend on time).

number of space steps	number of time steps	L^2 error	L^∞ error
80×80	5	0.0026	0.0062
80×80	10	0.0023	0.0052
80×80	80	0.0023	0.0053
number of space steps	number of time steps	L^2 error	L^∞ error
20×20	80	0.0136	0.0332
40×40	80	0.0051	0.0118
80×80	80	0.0023	0.0053

Table 2.4: Error with varying number of time steps (resp. space steps) for the Crank Nicholson Scheme.

6.3 Application

We apply the method to a financial example: we compute the price of a put option of strike $K = 1$ and maturity $T = 1$. In this model, X represents the price of the underlying of the put, and Y represents the price of the forward variance swap on the underlying X. Therefore, the terminal condition is

$$v(T, x, y) = (K - x)_+ \,.$$

Numerically we compute the price of the option for larges values of Y (i.e., $Y \simeq 3$), in order to use Neumann conditions for large Y. This approach is coherent with the value of interest which are typically for Y lower (or of the order of) unity. The result is shown in Fig. 2.3

Appendix

7 Resolution of infinite linear systems

Case of infinite 2d matrices. We say that the set of real numbers $A = (A_{(i,j),(k,\ell)})_{1\le i,j,k,\ell}$ is an *infinite 2d* matrix if $\{(k,\ell),\ A_{(i,j),(k,l)} \neq 0\}$ is finite $\forall i, j \geq 1$ (A is also an "infinite" tensor). If $X = (X_{i,j})_{i,j\ge 1}$ then we denote $(AX)_{i,j} = \sum_{k,\ell\ge 1} A_{(i,j),(k,\ell)} X_{k,\ell}$. We also denote $X \geq 0$ if $X_{i,j} \geq 0$, $\forall i, j \geq 1$.

The previous results can be easily generalized to infinite 2d matrices. We state here the results without proof.

Proposition 7.19. *Let $A = (A_{(i,j),(k,\ell)})_{1\le i,j,k,\ell}$ be an infinite 2d matrix such that*
(i) For all $i, j \geq 1$, $\exists \delta_{ij} \geq 0$, $A_{ii} = \delta_{ij} + \sum_{(k,\ell)\neq(i,j)} |A_{(i,j),(k,\ell)}|$,
(ii) $A_{(i,j),(k,\ell)} \leq 0\ \forall (i, j) \neq (k, \ell)$,
(iii) $\delta_{i1} > 0$, $\forall i \geq 1$,
(iv) $\exists C > 0$, $\forall i, j \geq 1$, $\sum_{(k,\ell)\neq(i,j)} |A_{(i,j),(k,\ell)}| \geq C\delta_{ij}$,
(v) $\forall i \geq 1$, $\forall j \geq 2$, if $\delta_{ij} = 0$ then $\exists q_{ij} > 0$ such that

$$(AX)_{ij} = q_{ij}(-X_{i,j-1} + 2X_{i,j} - X_{i,j+1}).$$

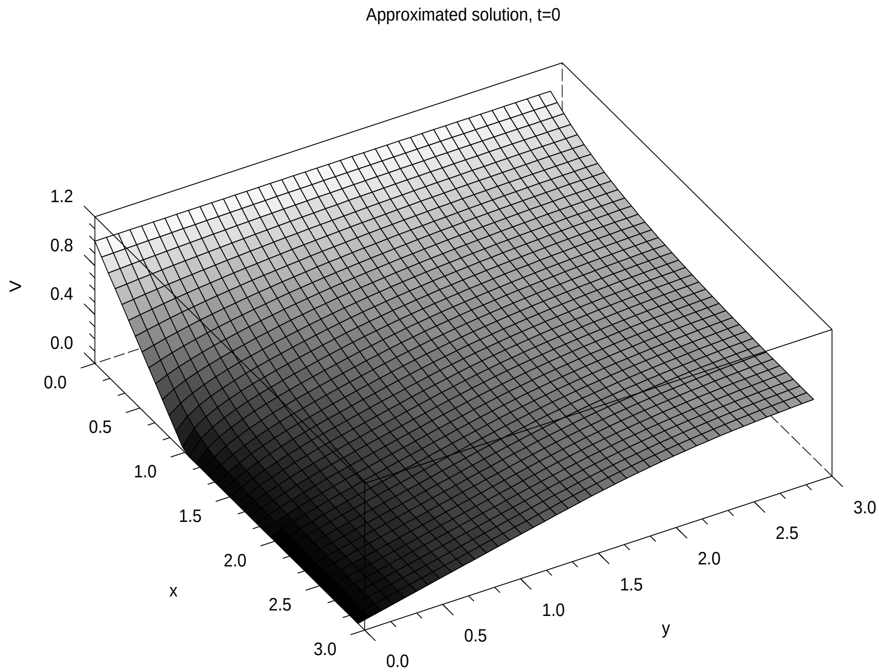


Figure 2.3: Surreplication price at time $t = 0$, with $T = 1$, $K = 1$ and payoff $(K - x)_+$.

1) Then A is monotone in the following sense: if $X = (X_{i,j})_{i,j\geq 1}$ is bounded from below and such that $\forall i, j \geq 1$, $\delta_{i,j} = 0 \Rightarrow (AX)_{i,j} = 0$, then

$$AX \geq 0 \Rightarrow X \geq 0.$$

2) If $b = (b_{ij})_{i,j\geq 1}$ is such that $\delta_{ij} = 0 \Rightarrow b_{i,j} = 0$, and $\displaystyle\max_{i,j\geq 1,\ \delta_{ij}>0} \frac{|b_{ij}|}{\delta_{ij}} < \infty$, then there is a unique X such that $AX = b$ and

$$\max_{i,j\geq 1} |X_{ij}| \leq \max_{i,j\geq 1,\ \delta_{ij}>0} \frac{|b_{ij}|}{\delta_{ij}}.$$

8 Properties of some infinite linear system

In this section we give some basic results for solving some specific infinite matrix system that are involved in our scheme.

Notations. We say that $A = (a_{ij})_{1\leq i,j}$, $i, j \in \mathbb{N}^*$, with $a_{ij} \in \mathbb{R}$ is an *infinite* matrix if $\{j \geq 1,\ a_{ij} \neq 0\}$ is finite $\forall i \geq 1$. If $X = (x_i)_{i\geq 1}$ then we denote $(AX)_i = \sum_{j\geq 1} a_{ij}x_j$. We also denote $X \geq 0$ if $x_i \geq 0$, $\forall i \geq 1$.

The following Lemma generalizes the monotony property of M-matrices.

Lemma 8.14 (monotony). *Let $A = (a_{ij})_{1\leq i,j}$ be a real infinite matrix such that*
(i) For all $i \geq 1$, $\exists \delta_i \geq 0$, $a_{ii} = \delta_i + \sum_{j\neq i} |a_{ij}|$,
(ii) $a_{ij} \leq 0$ $\forall i \neq j$,
(iii) $\delta_1 > 0$,
(iv) $\forall i \geq 1$, $\sum_j j\ a_{ij} \geq 0$.
(v) $\forall i \geq 2$, if $\delta_i = 0$ then $\exists q_i > 0$ such that $(AX)_i = q_i(-x_{i-1} + 2x_i - x_{i+1})$.
Then A is monotone in the following sense: if $X = (x_i)_{i\geq 1}$ is bounded from below and such that $\forall i \geq 1$, $\delta_i = 0 \Rightarrow (AX)_i = 0$, then

$$AX \geq 0 \Rightarrow X \geq 0.$$

Remark 8.1. *Note that from Lemma 8.14 we deduce the uniqueness of bounded solutions of $AX = b$ for any b such that $\delta_i = 0 \Rightarrow b_i = 0$.*

Proof of Lemma 8.14. Let $m = \min_{i\geq 1} x_i$.

Step 1. We first assume that there exists $i \geq 1$ such that $m = x_i$. Then

$$0 \leq a_{ii}x_i + \sum_{j\neq i} a_{ij}x_j = \delta_i x_i + \sum_{j\neq i} |a_{ij}|(x_i - x_j) \leq \delta_i x_i$$

If $\delta_i > 0$, then $x_i \geq 0$. In the case $\delta_i = 0$, by assumption (v) we obtain that $m = x_i = x_{i-1} = x_{i+1}$. In particular the minimum m is also reached by x_{i-1}. Since $\delta_1 > 0$, by a recursion argument we will arrive at a point j such that $\delta_j > 0$ and thus $x_j \geq 0$.

Step 2. In the general case we consider $Y = (y_i)$ with $y_i := x_i + \varepsilon\ i$ for some $\varepsilon > 0$. We note that $y_i \to +\infty$, hence $i \to y_i$ has a minimum. Also, $(AY)_i = (AX)_i + \varepsilon \sum_j j\ a_{ij} \geq 0$. Hence $AY \geq 0$ and $Y \geq 0$ by Step 1. Since this is true for any $\varepsilon > 0$, we conclude that $X \geq 0$. $\square$

Remark 8.2. *Note that in Lemma 8.14 we can relax the assumption (x_i) bounded from below by $\liminf_{i\to\infty} \frac{x_i}{i} \geq 0$.*

Proposition 8.20 (Existence of solutions for linear systems)**.** *We consider A, an infinite matrix, such that*
(i) $\forall i \geq 1$, $\exists \delta_i \geq 0$, $a_{ii} = \delta_i + \sum_{j\neq i} |a_{ij}|$,
(ii) $\delta_1 > 0$.
(iii) $\forall i \geq 2$, if $\delta_i = 0$ then $\exists q_i > 0$, such that $(AX)_i = q_i(-x_{i-1} + 2x_i - x_{i+1})$.
Let also $b = (b_i)_{i\geq 1}$ be such that

$$\forall i, \delta_i = 0 \Rightarrow b_i = 0, \qquad \text{and} \qquad \max_{k\geq 1,\ \delta_k\neq 0} \frac{|b_k|}{\delta_k} < \infty.$$

Then there exists a unique X, in the space of bounded sequences, such that $AX = b$, and furthermore we have

$$\max_{k\geq 1} |x_k| \leq \max_{k\geq 1,\ \delta_k\neq 0} \frac{|b_k|}{\delta_k}.$$

Proof. We look for solutions $x^{(n)} = (x_1^{(n)}, \dots, x_n^{(n)})^T \in \mathbb{R}^n$ of the first n linear equations of $AX = b$, and set also $x_k^{(n)} := 0$, $\forall k > n$. (Dirichlet type boundary conditions on the right border). This leads to solve the finite dimensional system

$$A^{(n)} x^{(n)} = b^{(n)} \tag{8.1}$$

where $A^{(n)} := (a_{ij})_{1\leq i,j\leq n}$ and $b^{(n)} := (b_1, \dots, b_n)^T$,

Lemma 8.15. *There exists a unique $x^{(n)}$ solution of* (8.1) *and furthermore it satisfies the inequality*

$$\max_{1\leq k\leq n} |x_k^{(n)}| \leq \max_{1\leq k\leq n,\ \delta_k\neq 0} \frac{|b_k|}{\delta_k}. \tag{8.2}$$

Proof of Lemma 8.15. Suppose that $x^{(n)}$ exists, and let i be such that $|x_i^{(n)}| = \max_{1\leq j\leq n} |x_j^{(n)}|$. Note that we still have $\forall 1 \leq i \leq n$, $a_{ii}^{(n)} = \delta_i + \sum_{j\neq i} |a_{ij}^{(n)}|$. If $\delta_i > 0$,

$$|b_i| \geq |a_{ii}^{(n)} x_i^{(n)}| - \sum_{j\neq i} |a_{ij}^{(n)}||x_j^{(n)}| \geq \delta_i |x_i^{(n)}|$$

thus $|x_i^{(n)}| \leq \frac{|b_i|}{\delta_i}$. If $\delta_i = 0$, we consider

$$i_0 := \sup\{k < i,\ \delta_k > 0\}.$$

(i_0 exists since $\delta_1 > 0$). Then $-x_{k-1}^{(n)} + 2x_k^{(n)} - x_{k+1}^{(n)} = b_k/q_k = 0$ for $k = i_0 + 1, \dots, i$, and $x_{k+1}^{(n)} - x_k^{(n)} = const = c_0$ for $k = i_0, \dots, i$. But $x_i^{(n)}$ is an extremum of $x_{i-1}^{(n)}$, $x_i^{(n)}$ and $x_{i+1}^{(n)}$. This implies that $x_{i-1}^{(n)} = x_i^{(n)} = x_{i+1}^{(n)}$, and thus $c_0 = 0$ and $x_{i_0}^{(n)} = x_i^{(n)}$ is also an extremum. Since $\delta_{i_0} > 0$, we can estimate $|x_{i_0}^{(n)}|$ as before. This implies the invertibility of $A^{(n)}$, and thus the uniqueness of $x^{(n)}$. $\square$

Now we shall prove that the sequence $X^{(n)} = (x^{(n)}, 0, 0, \dots)^T$, which satisfies already $||X^{(n)}||_\infty \leq C := \max_{\delta_k \neq 0} \frac{|b_k|}{\delta_k}$, converges pointwisely towards a solution X of the problem. We first suppose that $b \geq 0$. We can see that $A^{(n)}$ is still a monotone matrix (following the proof of Lemma 8.14). Hence $x^{(n)} \geq 0$. Now we consider $x^{(n+1)}$ and for $i \leq n$ we see that

$$\left(A^{(n)} x^{(n+1)}\right)_i = b_i - a_{i,n+1} x_{n+1}^{(n+1)} \geq b_i = \left(A^{(n)} x^{(n)}\right)_i.$$

Hence we obtain that

$$(x_1^{(n+1)}, \dots, x_n^{(n+1)})^T \geq (x_1^{(n)}, \dots, x_n^{(n)})^T,$$

and in particular $X^{(n)} \leq X^{(n+1)}$. Since $||X||_\infty \leq C$, we obtain the (pointwise) convergence of $X^{(n)}$ towards some vector X such that $||X||_\infty \leq C$. In the general case, we can decompose $b = b^+ - b^-$ with $b^+ = \max(b, 0)$, $b^- = \max(-b, 0)$, and proceed in the same way. We obtain the pointwise convergence of $X^{(n)} = X^{(n),+} - X^{(n),-}$ towards some X, with $X^{(n),\pm} \geq 0$ and $||X^{(n),\pm}||_\infty \leq C$, hence also $||X||_\infty \leq C$.

Since $\{j,\ a_{ij}^{(n)} \neq 0\}$ is finite, for any given i we can pass to the limit $n \to \infty$ in $\sum_{j \geq 1} a_{ij}^{(n)} x_j^{(n)} = b_i$, and obtain $(AX)_i = b_i$. □

Case of infinite 2d matrices. We say that the set of real numbers $A = (A_{(i,j),(k,\ell)})_{1 \leq i,j,k,\ell}$ is an *infinite 2d* matrix if $\{(k,\ell),\ A_{(i,j),(k,l)} \neq 0\}$ is finite $\forall i, j \geq 1$ (A is also an "infinite" tensor). If $X = (X_{i,j})_{i,j \geq 1}$ then we denote $(AX)_{i,j} = \sum_{k,\ell \geq 1} A_{(i,j),(k,\ell)} X_{k,\ell}$. We also denote $X \geq 0$ if $X_{i,j} \geq 0$, $\forall i, j$.

The previous results can be easily generalized to infinite 2d matrices. We state here the results without proof.

Proposition 8.21. *Let $A = (A_{(i,j),(k,\ell)})_{1 \leq i,j,k,\ell}$ be an infinite 2d matrix such that*
(i) For all $i, j \geq 1$, $A_{(i,j),(i,j)} = \delta_{ij} + \sum_{(k,\ell) \neq (i,j)} |A_{(i,j),(k,\ell)}|$ with $\delta_{ij} \geq 0$,
(ii) $A_{(i,j),(k,\ell)} \leq 0\ \forall (i,j) \neq (k,\ell)$,
(iii) $\delta_{i1} > 0$, $\forall i \geq 1$,
(iv) $\forall i, j \geq 1$, $\sum_{(k,\ell)} (k+\ell) A_{(i,j),(k,\ell)} \geq 0$,
(v) $\forall i \geq 1$, $\forall j \geq 2$, if $\delta_{ij} = 0$ then $\exists q_{ij} > 0$ such that

$$(AX)_{ij} = q_{ij}(-X_{i,j-1} + 2X_{i,j} - X_{i,j+1}).$$

1) Then A is monotone in the following sense: if $X = (X_{i,j})_{i,j \geq 1}$ is bounded from below and such that $\forall i, j \geq 1$, $\delta_{i,j} = 0 \Rightarrow (AX)_{i,j} = 0$, then

$$AX \geq 0 \Rightarrow X \geq 0.$$

2) If $b = (b_{ij})_{i,j \geq 1}$ is such that $\delta_{ij} = 0 \Rightarrow b_{i,j} = 0$, and $\max\limits_{i,j \geq 1,\ \delta_{ij} > 0} \dfrac{|b_{ij}|}{\delta_{ij}} < \infty$, then there is a unique bounded X such that $AX = b$, and furthermore

$$\max_{i,j \geq 1} |X_{ij}| \leq \max_{i,j \geq 1,\ \delta_{ij} > 0} \frac{|b_{ij}|}{\delta_{ij}}.$$

9 Convergence of the Howard algorithm

In this section we prove the following result.

Proposition 9.22. *Let S be a compact set, and $\mathcal{A} := S^{\mathbb{N}}$, the set of infinite sequences of S. For all $w \in \mathcal{A}$, let $A(w) := (a_{ij}(w))_{i,j\geq 1}$ be an infinite matrix, and $b(w) := (b_i(w))_{i\geq 1}$. We assume furthermore that*
(i) If $w = (w_i)_{i\geq 1}$, $a_{ij}(w)$ depends only of w_i, and also $b_i(w)$ depends only of w_i, and this dependence is continuous.
(ii) $\forall i$, $\sup_{w\in\mathcal{A}} \big(Card\{j,\ a_{ij}(w) \neq 0\}\big) < \infty$.
(iii) (monotony) For all $w \in \mathcal{A}$ and X bounded,

$$A(w)X \geq 0 \quad \Rightarrow \quad X \geq 0.$$

(iv) $\exists C \geq 0$, $\forall w \in \mathcal{A}$, $\exists X$ solution of $A(w)X = b(w)$ and such that

$$||X||_\infty \leq C.$$

Then
(i) there exists a unique bounded solution X to the problem

$$\min_{w\in\mathcal{A}}(A(w)X - b(w)) = 0. \tag{9.1}$$

(ii) the Howard algorithm as defined in section 4 converges pointwisely towards X.

Remark 9.1. *Proposition 9.22 can then be adapted in order to prove Proposition 4.18. The proof is left to the reader.*

Proof. Let us first check the uniqueness. Let X and Y be two solutions, and let $\bar{w}$ be an optimal control associated to Y. Then

$$\begin{aligned} A(\bar{w})Y - b(\bar{w}) &= 0 \\ &= \min_{w\in\mathcal{A}}(A(w)X - b(w)) \\ &\leq A(\bar{w})X - b(\bar{w}). \end{aligned}$$

Hence $A(\bar{w})(Y - X) \leq 0$ and thus $Y \leq X$ using the monotony property. We can prove $Y \geq X$ in the same way, hence $X = Y$ which proves uniqueness.

The existence now is obtained by considering the sequence X^k and controls w^k as in the Howard algorithm of section 4.

We first remark that for all $k \geq 0$, $X^k \leq X^{k+1}$, because

$$\begin{aligned} A(w^{k+1})X^{k+1} - b(w^{k+1}) &= 0 \\ &= A(w^k)X^k - b(w^k) \\ &\geq \min_{w}(A(w)X^k - b(w)) \\ &\geq A(w^{k+1})X^k - b(w^{k+1}) \end{aligned}$$

and using the monotony of $A(w^{k+1})$. Also, X^k is bounded. Hence X^k converges pointwisely towards some bounded X. It remains to show that X satisfies (9.1).

Let $F_i(X)$ be the i-th component of $\min_{w\in\mathcal{A}}(A(w)X - b(w))$, i.e.

$$F_i(X) = \min_{w\in\mathcal{A}}(A(w)X - b(w))_i$$

For a given i, since $(A(w)X)_i$ involves only a finite number of matrix continuous coefficients $(a_{ij}(w))_{j\leq j_{\max}}$, we obtain that $\lim_{k\to\infty} F_i(X^k) = F_i(X)$. Also by compactness of S, by a diagonal extraction argument, there exists a subsequence of $(w^k)_{k\geq 0}$, denoted w^{ϕ_k}, that converges pointwisely towards some $w \in \mathcal{A}$.

Passing to the limit in $(A(w^{\phi_k})X^{\phi_k} - b(w^{\phi_k}))_i = 0$, we obtain $(A(w)X - b(w))_i = 0$. On the other hand,

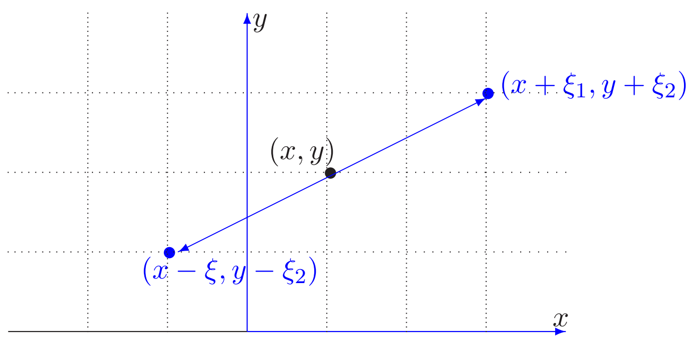
$$\begin{aligned} F_i(X) &= \lim_{k\to\infty} F_i(X^{\phi_k - 1}) \\ &= \lim_{k\to\infty} \Big(A(w^{\phi_k})X^{\phi_k} - b(w^{\phi_k})\Big)_i \\ &= \big(A(w)X - b(w)\big)_i \end{aligned}$$

Hence $F_i(X) = 0$, $\forall i$, which concludes the proof. □

10 Points on the boundary

We present in this section another way to consider the point near to the boundary in the discretization of the second order term. Consider the grid points that are close to the boundaries $x = 0$, $y = 0$. Fixes an order p_{max}, for theses points, the discretization of the second order term could involve some author points which are out of the grid.

Then we modify the expression of the elementary diffusion. Let us explain this modification on a simple example drawn in the following picture

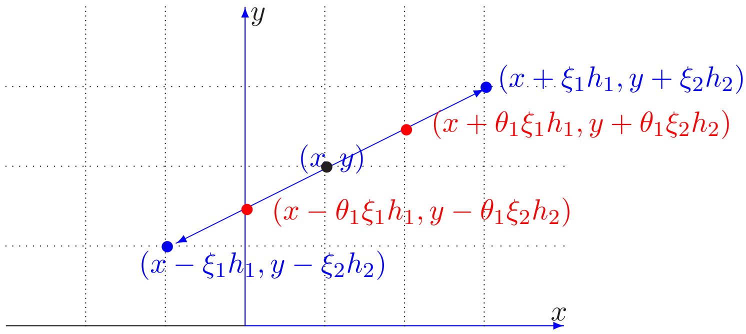


The direction of the diffusion (the vector $\longrightarrow$) points toward a grid points $(x \pm \xi_1 h_1, y \pm \xi_2 h_2)$ in the neighborhood of order 2. However, $(x - \xi_1 h_1, y - \xi_2 h_2)$ is out of the grid, which is delimited by the positive part of the x-axis and the positive part of the y-axis.

Introduce a parameter $\theta_1 \in [0, 1]$ and the associated point $(x - \theta_1\xi_1 h_1, y - \theta_1\xi_2 h_2)$. The real θ_1 is chosen in such a way that the point $(x - \theta_1\xi_1 h_1, y - \theta_1\xi_2 h_2)$ is on the diffusion direction and belonging to the y-axis (the intersection between the y-axis and the vector formed by (x, y) and $(x - \xi_1 h_1, y - \xi_2 h_2)$). Although $(x - \theta_1\xi_1 h_1, y - \theta_1\xi_2 h_2)$ is not a grid point, we will use it in the scheme, because the function v_h is known on the axis $x = 0$. The elementary diffusion becomes

$$\Delta_{\xi,\theta}\phi(x,y) = \frac{\phi(x+\theta_1\xi_1 h_1, y+\theta_1\xi_2 h_2) + \phi(x-\theta_1\xi_1 h_1, y-\theta_1\xi_2 h_2) - \phi(x,y)}{\theta_1^2}, \qquad (10.1)$$

where $\theta_1 \in [0,1]$ is chosen such that $(x-\theta_1\xi_1 h_1, y-\theta_1\xi_2 h_2)$ and $(x+\theta_1\xi_1 h_1, y+\theta_1\xi_2 h_2)$ are in the domain $[0,+\infty)^2$.



In general case, for ξ is in the stencil $\mathcal{S}_{p_{\max}}$ and (x,y) in the grid G_h, if the points $(x \pm \xi_1 h_1, y \pm \xi_2 h_2)$ should be used in the approximation of the covariance matrix and if thery are out of the domain $[0,+\infty)^2$, then we modify the elementary diffusion Δ_ξ by:

$$\Delta_{\xi,\theta}\phi(x,y) = \frac{\theta_2\phi(x+\theta_1\xi_1, y+\theta_1\xi_2) + \theta_1\phi(x-\theta_2\xi_1, y-\theta_2\xi_2) - 2(\theta_1+\theta_2)\phi(x,y)}{\theta_1\theta_2(\theta_1+\theta_2)},$$

where $\theta_1, \theta_2 \in [0,1]$ are such that $(x+\theta_1\xi_1 h_1, y+\theta_1\xi_2 h_2)$ and $(x-\theta_2\xi_1 h_1, y-\theta_2\xi_2 h_2)$ are in the domain.

Therefore, the scheme (3.4) should be written:

$$v_h(T,x,y) = g(x) = v_h(t,x,0), \quad v_h(t,0,y) = g(0), \qquad (10.2)$$

$$\min_{\alpha_1^2+\alpha_2^2=1} \{-\alpha_1^2\delta_t v_h(t,x,y) - \alpha_1^2\mu\delta_y v_h(t,x,y) - \frac{1}{2}\sum_{\xi\in\mathcal{S}} \gamma_\xi^{\alpha_1,\alpha_2}\Delta_{\xi,\theta} v_h(t,x,y)\} = 0, \qquad (10.3)$$

for $t < T - \Delta t$.

This scheme satisfies consistency property, in the sense of Proposition 3.17.

Part II

An uncertain volatility model

Chapter 3

Option pricing with uncertain volatility and tolerance against losses.

This chapter studies European options hedging and pricing with volatility risk in an uncertain volatility model as in [3]. It is well known that the superreplication criterion leads, when the volatility is not a priori bounded, to trivial strategies and too expensive costs. We propose a criterion, which does not require to hedge almost surely the option's payoff, as we admit some possible losses depending on the realized volatility until maturity. These acceptable losses are described through a function depending on time, asset prices and volatility. We show that our framework recovers solutions obtained in uncertain volatility model [3], and gamma constraints [65].

Key words: uncertain volatility, super replication, viscosity solutions.

1 Introduction

An important part of the literature about option pricing is motivated by the fact that the Black and Scholes model fails to capture precisely the behavior of the financial asset prices. Indeed, two important observations contradict that model. The first one is the discontinuity of asset prices through time (i.e. prices jumps), and the second one is the heteroskedasticity of asset returns, that leads to study models which features random volatility. In this chapter, we are interested in the consequences of stochastic volatility on option pricing. To formalize this problem, there are two possible approaches. The first one is to consider that, despite the stochastic volatility, the market is still complete, that is, one can perfectly hedge an option with the underlying and, say, another option. Then the problem is reduced to a calibration problem. But this assumption is valid only if there exists a liquid market for vanilla options. This is the case for large indexes and stocks but it is generally false for more exotic underlyings like investment funds. The second approach is to consider that the market is incomplete, but then the question of the pricing criterion arises. Many can be found in the literature: for instance mean-variance hedging, see the surveys by Pham [57] and [62], and indifference pricing (see [61]). But in order to use these techniques, one must know precisely the dynamics of the underlying and its volatility. Indeed, the effects of model misspecification in those frameworks are not well known. Also there may not be enough available data to estimate the corresponding models, with a satisfying confidence interval. But this problem can be circumvented if one prices and hedges the option with a robust criterion. In this case, one would take into account model uncertainty. This kind of uncertainty has been studied, for instance, recently in [25]. The advantage of that kind of methods, is that one does not need to know precisely the dynamic properties of the underlying, but rather some knowledge or priors about the probability distribution of the volatility. The first step was the study of the robustness of the Black and Scholes formula in [35]. Then, in [3] and [40], the authors derived a pricing formula that enables to perform a superhedge of the option as long as the volatility evolves inside a given interval. The problem of that method is that one does not control what happens if the volatility goes out of that interval, and an inappropriate interval may theoretically leads to bankruptcy. Hence one may be tempted to consider a very large interval, but this would lead to huge selling prices. In our work we consider a criterion that enables to control the losses of the option's seller in any case, without necessarily involving superreplication. This criterion might prove useful to control volatility risk in all possible outcomes with a reasonable price, and without a precise knowledge of the volatility dynamics. We will use the same kind of model as in [3], that is, an uncertain volatility model. It is modeled here as an adverse control, that is, we consider a worst case scenario. Therefore, we will not postulate the dynamics of the volatility. However, in contrast [3], the volatility σ will not be constrained to lie in a bounded set. Instead, all the assumptions concerning the volatility will be contained in an admissible losses function. In practice, the option seller should admit to lose money for levels of realized volatility that he thinks to be unprobable. Indeed, admitting losses has the effects of reducing the options price, hence increasing the margins and profits in the more probable outcomes. As a limit case, if he is absolutely certain that the volatility can never take values above a given level, he should accept to risk infinite losses the volatility effectively takes values above that threshold.

This function will modify the payoff of the option we want to price, and then we will use a more classical super-replication criterion to price this modified payoff. The aim of this chapter is to characterize that price as the unique viscosity solution of a PDE. There are two major difficulties, though. First of all, we do not have any dynamic programming principle for the super-replication problem, as our framework differs from [32] because the volatility is not a priori bounded. This difficulty is circumvented by introducing a dual representation of the problem which is an expectation maximization problem. The second difficulty is to obtain a verification theorem for the solution of a PDE which is not regular enough to apply Itô's formula. This is why we need to introduce some regularization techniques.

This chapter is organized as follows: First, we introduce our model and some technical assumptions, and define the option seller's price. Then, we state the main results and derive the viscosity property for the dual representation. In the next section, we identify the original problem with its dual representation. Finally, we prove the comparison principle to obtain uniqueness of the viscosity solution of the equation satisfied by the price.

2 Problem formulation

2.1 Description

We consider a market with $d+1$ asset. There is a riskless asset which will be taken as the numeraire, so we can suppose the interest rate equal to 0. For the d risky assets, we consider a d-dimensional uncertain volatility model:

$$dS_t = \operatorname{diag}(S_t)\sigma_t dW_t \tag{2.1}$$

where W_t is a d-dimensional Brownian motion, and σ is a d-dimensional positive symmetric matrix valued process. The set Σ of admissible processes $\sigma = (\sigma_t)_{t\in[0,T]}$ will be defined later. Working with a multidimensional model is obviously useful when considering covariance risk for basket or exchange options on several underlings. But it also is possible to interpret one of the prices as the underlying of an option that we are trying to sell, and the other options as liquid options on the same underlying. Then we can obtain a robust version of an existing volatility model.

The agent has a tolerance to losses which is a function of the prices, the date, and more crucially of the volatility. This is modelled through a function

$$f : [0,T] \times \mathbb{R}^d \times \mathcal{S}_d \to \mathbb{R} \cup \{+\infty\}$$

where $f(t, S_t, \sigma_t^2)$ represents the maximum loss that the option's seller admits per unit of time at time time t. Now, we define the lowest selling price of the agent as:

$$v(t,s) = \inf_z \Big\{ z \in \mathbb{R} : \text{ there exists } \pi \in \mathcal{A}_1 \text{ such that} \tag{2.2}$$

$$z + \int_t^T \pi_u dS_u^{t,s,\sigma} \geq g(S_T^{t,s,\sigma}) - \int_t^T f(u, S_u^{t,s,\sigma}, \sigma_u^2) du \text{ a.s. for all } \sigma \in \Sigma \Big\}$$

Here, $\mathcal{A}_1$ is the set of adapted processes such that $\int_0^T \pi_t dS_u^{0,s,\sigma}$ is almost surely bounded from below for all $\sigma \in \Sigma$, and Σ is defined as:

$$\Sigma = \Big\{ \sigma = (\sigma_t)_{t\in[0,T]} \text{ a.s. bounded and adapted processes valued in } \mathcal{S}_+^d$$

$$\text{s.t. } \int_0^T f(t, S_t^{0,s,\sigma}, \sigma_t) dt \text{ is a.s. bounded } \Big\},$$

the set of adapted processes taking values in $\mathcal{S}_+^d$ that are a.s. bounded. $S_\cdot^{t,s,\sigma}$ is the solution of (2.1), for $u \geq t$, under volatility $\sigma \in \Sigma$ and starting from $S_t = s$ at time t. The super-replication cost corresponds to the particular case $f \equiv 0$. Unfortunately, there is no existing dynamic programming principle for this problem. For instance, one cannot use the results of [32] as the volatility does not admit a uniform bound. Following the dual approach to superreplication criterion (see [36]), we introduce the classical stochastic control problem:

$$w(t,s) = \sup_{\sigma \in \Sigma} \mathbb{E}[g(S_T^{t,s,\sigma}) - \int_t^T f(u, S_u^{t,s,\sigma}, \sigma_u^2)du] \tag{2.3}$$

With this formulation, one can use the classical results of stochastic control to derive the Hamilton-Jacobi-Bellmann equation. Then, in section 4, we will prove by a hedging argument that problems (2.3) and (2.2) admit the same solution, that is, $v = w$. The only technical difficulty is that, as usual, w is not supposed to be regular, and we must use regular approximations to apply Ito's formula to derive a hedging argument. An advantage of that proof, is that we will be able to show that, at least with approximated prices, the optimal hedging is a classical delta hedging. In other words, the hedging portfolio is, in quantity, the gradient of the price w.r.t. the spot price, as in the usual Black-Scholes framework. It is interesting since it shows that one does not need to know the actual volatility of the underlying to perform the hedge. This might be useful when that volatility is difficult to measure precisely. We end this section by stating some technical conditions.
Here, we state an assumption which will prove useful to simplify some demonstrations:

Assumption 2.1. *For each $(t,s) \in [0,T] \times \mathbb{R}^d$, the function:*

$$\begin{aligned} f(t,s,.) : \mathcal{S}_d &\to \mathbb{R} \\ \sigma^2 &\to f(t,s,\sigma^2) \end{aligned}$$

is convex and lower semicontinuous.

This is not a restrictive assumption, as we will see that the pricing PDE only involves the Fenchel transform of f with respect to σ^2. Therefore the resulting price can be as well associated to the convex lower semicontinuous envelope of f, if it is not already convex l.s.c. To ensure the viscosity property, we need the following assumption, in order to control the dependency of f w.r.t. time and spot prices:

Assumption 2.2. *There exists a positive function:*

$$h : (0,+\infty) \to (0,+\infty)$$

such that, for any $\varepsilon > 0$, and $(t,s,\sigma^2) \in [O,T] \times (0,+\infty)^d \times \mathcal{S}_+^d$

$$|f(t,s,\sigma^2) - f(t',s',\sigma^2)| \leq \varepsilon \text{ if } |t-t'| \leq h(\varepsilon) \text{ and } \left\|s-s'\right\|_\infty \leq h(\varepsilon)$$

We introduce another assumption, in order to make sure that the limit price of the option when time approaches maturity is above the exercise price. In other words, we make sure that we do not tolerate too much losses, which may introduce arbitrage opportunities for other agents.

Assumption 2.3. *The function f is bounded from below, and there exists a bounded continuous feedback control*

$$\begin{aligned} \sigma : (0,+\infty)^d &\to \mathcal{S}_+^d \\ (t,s) &\to \sigma(t,s) \end{aligned}$$

and a constant C such that $f(t,s,\sigma^2(t,s)) < C$, for all $(t,s) \in [O,T] \times (0,+\infty)$.

Therefore, we will only consider acceptable losses functions such that there always exists a bounded volatility scenario for which losses will be bounded. Now we introduce an assumption on the Fenchel transform $\tilde{f}$ of f (see (3.2)) which will ensure the uniqueness result for the PDE satisfied by the value function. To fully understand it, one must recall that under assumption 2.2, the domain of f w.r.t. σ^2 is independent of t, s.

Assumption 2.4. *For any $\varepsilon > 0$,there exists K_ε such that:*

$$\tilde{f} \text{ is } K_\varepsilon - Lip \text{ in } int_\varepsilon(dom\left(\tilde{f}(t,s,.)\right))\forall(t,s)$$

and $0 \in int\left(dom(\tilde{f}(t,s,.))\right)\forall(t,s)$

Where the ε-interior a subset A of $\mathcal{M}^d(\mathbb{R})$ is defined as:

$$int_\varepsilon(A) = \{x \in A \text{ such that } \mathcal{B}(x,\varepsilon) \subset A\}$$

and $\mathcal{B}(x,\varepsilon)$ is the ball of radius ε centered on x. This assumption is verified for example if $dom(f(t,s,.))$ is bounded uniformly in (t,s). Indeed, in this case, $dom(\tilde{f}(t,s,.)) = \mathcal{S}_d$ and $\tilde{f}(t,s,.)$ is Lipschitz continuous uniformly in (t,s). A typical example where this assumption does not hold is, in dimension 1, $f(t,s,\sigma^2) = \sigma^4$. Indeed, this implies a quadratic behavior for both f and $\tilde{f}$, which contradicts the assumption.
It is interesting to see that, in dimension 1, this assumption only implies that either f or $\tilde{f}$ are uniformly Lipschitz continuous on an interval of type $\sigma^2 \in [C,+\infty)$, and that the upper semicontinuous envelope of f, f^*, is not uniformly equal to $+\infty$. This is indeed a much simpler statement. Sadly, this simple formulation does not seem to hold in dimension higher than 1. The last assumption concerns the payoff function, and is used to prove uniqueness of the solution of the characteristic PDE.

Assumption 2.5. *The payoff g is Lipschitz continuous and bounded.*

3 The PDE representation

3.1 Operator

As it is a classical problem in optimal control theory, the first candidate equation verified by the value function of the dual control problem (2.3) is:

$$-\frac{\partial v}{\partial t} - \tilde{f}(t,s,\mathrm{diag}\,[s]\,D^2 v \mathrm{diag}\,[s]) = 0 \tag{3.1}$$

where

$$\tilde{f}(t,s,A) = \sup_{\sigma^2} \left\{ \frac{1}{2} Tr\left(A\sigma^2\right) - f(t,s,\sigma^2) \right\} \tag{3.2}$$

is the convex conjugate of f. However,while this is the correct equation when $\tilde{f}$ is finite, we should take into account in our context that $\tilde{f}$ may take infinite values and so adjust the above PDE. This is why we introduce the second operator G. We know, from the definition of $\tilde{f}$, that the domain of $\tilde{f}$ is convex. This leads us to introduce the signed distance to the complementary of the domain of $\tilde{f}$:

$$\hat{G}(t,s,A) = \begin{cases} \inf\left\{ |B| \text{ s.t. } A+B \notin dom(\tilde{f}(t,s,.)) \right\} \text{ if } A \in dom(\tilde{f}(t,s,.)) \\ -\inf\left\{ |B| \text{ s.t. } A+B \in dom(\tilde{f}(t,s,.)) \right\} \text{ if } A \notin dom(\tilde{f}(t,s,.)) \end{cases}$$

With assumption 2.2, one can prove easily that $dom(\tilde{f}(t,s,.))$ does not depend of t and s. Hence we will denote it $dom(\tilde{f})$. Therefore, $\hat{G}(t,s,.)$ does not depends of (t,s). We immediatly see that $\tilde{f}(t,s,.)$ is increasing, that is $\tilde{f}(A) \geq \tilde{f}(B)$ for any $A \geq B$. Hence, if $B \notin dom(\tilde{f})$ then $A \notin dom(\tilde{f})$. Therefore, we can conclude that G is decreasing. In the next lemma, we prove that this function is also Lipschitz and concave.

Lemma 3.1. *With assumption 2.2, for any $A \in \mathcal{S}_d$, one has:*

$$\hat{G}(t,s,A) = \hat{G}(t',s',A) = G(A)$$

for all $(t,t',s,s') \in [0,T]^2 \times \mathbb{R}^{2d}$ Furthermore, the function, independent of (t,s)

$$\begin{aligned} G : \mathcal{S}_d &\to \mathbb{R} \\ A &\to \hat{G}(t,s,A) \end{aligned}$$

is Lipschitz continuous and concave.

Proof. For any $A' \geq A$, if $A+B \notin dom(\tilde{f})$, then $A'+B \notin dom(\tilde{f})$. Hence $G(A') \leq G(A)$. But, on the other hand, if $A'+B \notin dom(\tilde{f})$, then $A+(A'-A)+B \notin dom(\tilde{f})$, hence $G(A') + |A'-A| \geq G(A)$. As G is Lipschitz of constant 1 for any $A' \geq A$, then it Lipschitz is on the whole space. Indeed, any symmetric matrix can be decomposed into a positive and a negative part. □

Example 3.1. *A very interesting case, in the one-dimensional framework, is introduced in [63]. It leads to an equation one can obtain using the acceptable loss function:*

$$f(t,s,\sigma) = \begin{cases} +\infty \text{ } if \text{ } \sigma < \hat{\sigma} \\ \frac{1}{2}\Gamma_+(\sigma^2 - \hat{\sigma}^2) \text{ } if \text{ } \sigma \geq \hat{\sigma} \end{cases}$$

Wher $\hat{\sigma}$ and Γ_+ are given strictly positive constants. The conjugate of the acceptable loss function is therefore:

$$\tilde{f}(t,s,s^2\frac{\partial^2 v}{\partial s^2}) = \begin{cases} +\infty \text{ } if \text{ } \frac{1}{2}s^2\frac{\partial^2 v}{\partial s^2} > \Gamma_+ \\ \frac{1}{2}s^2\sigma^2\frac{\partial^2 v}{\partial s^2} \text{ } if \text{ } \frac{1}{2}s^2\frac{\partial^2 v}{\partial s^2} \leq \Gamma_+ \end{cases}$$

The authors of [63] obtain the proper equation, where $\tilde{f}$ is bound to take finite values:

$$\min\left\{-\frac{\partial v}{\partial t} - \tilde{f}\left(t, s, s^2 \min\left\{\frac{\partial^2 v}{\partial S^2}, \Gamma^+\right\}\right), \Gamma_+ - \frac{1}{2}S^2\frac{\partial^2 v}{\partial S^2}\right\} = 0 \tag{3.3}$$

This equation is, in this case, formally equivalent to (3.1). This form has the advantage to introduce only Lipschitz operators, which is very useful to prove uniqueness of its solutions. In our work, we try to obtain such a smooth formulation, but the general form of $\tilde{f}$ introduces new difficulties, especially in the multidimensional case.

Indeed, we shall justify later that, in order to obtain a regular operator, a suitable formulation of the PDE is:

$$\begin{aligned} F(t, s, \frac{\partial v}{\partial t}, D_s^2 v) = \sup_{A \geq 0} \Big\{ \min \Big\{ & -\frac{\partial v}{\partial t} - \tilde{f}\left(t, s, \operatorname{diag}[s] D^2 v \operatorname{diag}[s] - A\right) \\ & , \mathbf{1}_{A=0} G\left(\operatorname{diag}[s] D^2 v \operatorname{diag}[s]\right) - tr(A) \Big\}\Big\} = 0 \end{aligned} \tag{3.4}$$

3.2 Main result

The main result of the chapter is the characterisation of the price through equation (3.4), together with a terminal condition.

Theorem 3.2. *Let assumptions 2.2, 2.3, 2.4 and 2.5 hold. Then v is continuous and is the unique bounded viscosity solution of equation (3.4) with terminal condition $v(T^-, .) = \hat{g}$, where $\hat{g}$ is characterised as the unique bounded viscosity solution of:*

$$\min\left\{\hat{g}(s) - g(s), G(diag[s] D^2 \hat{g} diag[s])\right\} = 0 \tag{3.5}$$

Proof. With assumptions 2.2 and 2.3, the viscosity property of the dual control problem value function w is proved by propositions 3.1 and 3.2. The viscosity property is given for the terminal condition $\hat{g}$ of w by proposition 3.3. With assumptions 2.3 and 2.5, it is easily proved that the exists a constant C such that:

$$\hat{g} - C < w < \hat{g} + C$$

Hence, uniqueness of the solution of (3.4) is ensured by proposition 5.1. Finally, equality between v and w is treated in section 4. Uniqueness and continuity of bounded viscosity solutions of equation (3.5) can be proved by similar but simpler arguments, which are not exposed in this work for the sake of conciseness. $\square$

The next paragraphs relate rigorously, by means of viscosity solutions, the value function of control problem (2.3) to the variational PDE. We also characterize the corresponding terminal condition associated to (2.3). This is based on the following principle.

3.3 Viscosity solution property

Dynamic programming principle

Here, we state the classical dynamic programming principle (DPP) related to problem (2.3). It is the essential tool to prove that the value function w verifies equation (3.4). It is

discuted, among others, in [38] . It is separated in two parts, and is mathematically stated as follows:
(DP1) For all $\sigma \in \Sigma$ and $\theta \in \mathcal{T}_{t,T}$, set of stopping times valued in $[t,T]$:

$$w(t,s) \geq \mathbb{E}\left[-\int_t^\theta f(u, S_u^{t,s,\sigma}, \sigma_u^2)du + w(\theta, S_\theta^{t,x})\right] \tag{3.6}$$

(DP2) For all $\varepsilon > 0$, there exists $\hat{\sigma}^\varepsilon \in \Sigma$ s.t. for all $\theta \in \mathcal{T}_{t,T}$:

$$w(t,s) - \varepsilon \leq \mathbb{E}\left[-\int_t^\theta f(u, S_u^{t,s,\hat{\sigma}^\varepsilon}, (\hat{\sigma}_u^\varepsilon)^2)du + w(\theta, S_\theta^{t,x})\right] \tag{3.7}$$

Supersolution property

Proposition 3.1. *Assume assumption 2.2. Then, w is a viscosity supersolution of the HJB equation:*

$$-\frac{\partial w}{\partial t} - \tilde{f}\left(diag\,[s]\, D^2 w diag\,[s]\right) = 0,\ \ (t,s) \in [0,T) \times \mathbb{R} \tag{3.8}$$

Proof. This proof is classical in stochastic control theory, see for example [59] or [58] for details. The only difficulty here is that f can take infinite values, but it can be bypassed with assumption 2.2. Let $(\bar{t}, \bar{s}) \in [0,T) \times \mathbb{R}_+$ and $\varphi \in C^2\left([0,T) \times \mathbb{R}_+\right)$ a smooth test function satisfying:

$$0 = (w_* - \varphi)(\bar{t}, \bar{s}) = \min_{(t,s) \in [0,T) \times \mathbb{R}_+} (w_* - \varphi)(t,s) \tag{3.9}$$

By the definition of $\tilde{f}$, one has to prove that:

$$-\frac{\partial \varphi}{\partial t}(\bar{t}, \bar{s}) - \max_{\sigma^2 \in \mathcal{S}_d} \left\{\frac{1}{2} Tr\left(\sigma^2 \mathrm{diag}(\bar{s}) D^2 \varphi(\bar{t}, \bar{s}) \mathrm{diag}(\bar{s})\right) - f(\bar{t}, \bar{s}, \sigma^2)\right\} \geq 0$$

Let $\sigma^2 \in \mathbb{R}_+$. If $\sigma^2 \notin dom\left(f\left(\bar{t}, \bar{s}, .\right)\right)$ then, trivially:

$$-\frac{\partial \varphi}{\partial t}(\bar{t}, \bar{s}) - \frac{1}{2} Tr\left(\sigma_0^2 \mathrm{diag}(\bar{s}) D^2 \varphi(\bar{t}, \bar{s}) \mathrm{diag}(\bar{s})\right) + f(\bar{t}, \bar{s}, \sigma_0^2) = +\infty \geq 0$$

Now, if $\sigma_0^2 \in dom\left(f\left(\bar{t}, \bar{s}, .\right)\right)$, then with assumption 2.2 one can deduce easily that:

$$(t,s) \to f(t,s,\sigma_0^2) \in C^0\left([0,T) \times \mathbb{R}^+\right) \tag{3.10}$$

Now, we can use the standard arguments to prove the proposition. By definition of $w_*(\bar{t}, \bar{s})$, there exists a sequence (t_m, s_m) in $[0,T) \times \mathbb{R}^+$ such that:

$$(t_m, s_m) \to (\bar{t}, \bar{s}) \text{ and } w(t_m, s_m) \to w_*(\bar{t}, \bar{s})$$

when m goes to infinity. Then by continuity of φ and 3.9, we have:

$$\gamma_m := w(t_m, s_m) - \varphi(t_m, s_m) \to 0$$

when m goes to infinity. Let S_u^m the process starting from s_m at t_m and controlled by σ_0^2. Consider τ_m the first exit time of S^{t_m,s_m} from the open ball $B_\eta(s_m)$ with $\eta > 0$, and (h_m) a positive sequence such that:

$$h_m \to 0 \text{ and } \frac{\gamma_m}{h_m} \to 0$$

when m goes to infinity. Applying the first part of the dynamic programming principle DP1 (3.6) to $w(t_m, s_m)$ and $\theta_m = \tau_m \wedge (t_m + h_m)$, then using 3.9 and finally applying Ito's formula to φ between t_m and θ_m we get:

$$\frac{\gamma_m}{h_m} + E\left[\frac{1}{h_m}\int_{t_m}^{\theta_m} -\left(\frac{\partial \varphi}{\partial t} + \frac{1}{2}Tr\left(\sigma_0^2 \text{diag}(S_u^m) D^2\varphi(u, S_u^m)\text{diag}(S_u^m)\right)\right.\right.$$
$$\left.\left. + f\left(u, S_u^m, \sigma_0^2\right)\right) du\right] \geq 0$$

By the almost sure continuity of the trajectory S^m, then for m sufficiently large ($m > N(\omega)$), $\theta_m(\omega) = t_m + h_m$ a.s. Remarking that with property (3.10), the process inside the expectation has continuous trajectories and it is bounded independently of m. Hence we can use the mean-value theorem to find that the variable under the expectation converges a.s. to $-\mathcal{L}^{\sigma_0}\varphi(\bar{t}, \bar{s}) + f\left(\bar{t}, \bar{s}, \sigma_0^2\right)$. Finally, with the uniform bound, we can apply Lebesgue's dominated convergence theorem to obtain:

$$-\frac{\partial \varphi}{\partial t}(\bar{t}, \bar{s}) - \frac{1}{2}Tr\left(\sigma_0^2 \text{diag}(\bar{s}) D^2\varphi(\bar{t}, \bar{s})\text{diag}(\bar{s})\right) + f(\bar{t}, \bar{s}, \sigma_0^2) \geq 0$$

And we can conclude, due to the arbitrariness of σ_0. □

Subsolution property

Proposition 3.2. *Assume assumption 2.2. Then w is a viscosity subsolution of the HJB variational inequality:*

$$F(t, s, \frac{\partial w}{\partial t}, D_s^2 w) = \sup_{A \geq 0}\left\{\min\left\{-\frac{\partial w}{\partial t} - \tilde{f}\left(t, s, diag[s] D_s^2 w\, diag[s] - A\right)\right.\right. \tag{3.11}$$
$$\left.\left. , \mathbf{1}_{A=0} G\left(diag[s] D_s^2 w\, diag[s]\right) - tr\left(A\right)\right\}\right\} = 0 \tag{3.12}$$

Proof. As in [58], one has to introduce, for a given smooth function φ, the subset of $[0, T] \times \mathbb{R}_+^d$:

$$\mathcal{M}(\varphi) = \left\{(t, s) \in [0, T] \times \mathbb{R}_+^d : -\frac{\partial \varphi}{\partial t} - \tilde{f}\left(t, s, \text{diag}\left[s\right] D^2\varphi \text{diag}\left[s\right]\right) > 0\right.$$
$$\left. \text{and } G\left(t, s, \text{diag}\left[s\right] D^2\varphi(t, s)\text{diag}\left[s\right]\right) > 0\right\}$$

Then, we see that $(t, s) \in \mathcal{M}(\varphi)$ if and only if:

$$F(t, s, \frac{\partial \varphi}{\partial t}, D_s^2\varphi) > 0 \tag{3.13}$$

Indeed, if $F > 0$ and the supremum is attained for $A \neq 0$ in the definition of F, then we have $F \leq 0$ and we have a contradiction. If it is attained for $A = 0$, then

$$F(t, s, \frac{\partial \varphi}{\partial t}, D_s^2\varphi) = \min\left\{-\frac{\partial \varphi}{\partial t} - \tilde{f}\left(\text{diag}[s](D_s^2\varphi)\text{diag}[s]\right), G\left(\text{diag}[s] D_s^2\varphi \text{diag}[s]\right)\right)\right\}$$

which is strictly positive if and only if $(t, s) \in \mathcal{M}$. From this point we can exactly follow the demontration in [58], using (DP2) as main argument. □

3.4 Terminal condition

To define the terminal condition, we introduce the classical limit functions:

$$\underline{w}(s) = \liminf_{t \nearrow T, s' \to s} w(t,s') \text{ and } \overline{w}(s) = \limsup_{t \nearrow T, s' \to s} w(t,s')$$

By definition, $\underline{w} \leq \overline{w}$, $\underline{w}$ is l.s.c, and $\underline{w}$ is u.s.c. Now, we can characterize the terminal condition as follows:

Proposition 3.3. *Assume assumption 2.2, and that g is lower-bounded or satisfy a linear growth condition. If g is lower semicontinuous, then $\underline{w}$ is a viscosity supersolution of*

$$\min\left[\underline{w}(s) - g(s), G\left(T, s, diag\left[s\right] D^2 \underline{w}(s) diag\left[s\right]\right)\right]$$

If g is upper semicontinuous, then $\overline{w}$ is a viscosity subsolution of

$$\min\left[\overline{w}(s) - g(s), G\left(T, s, diag\left[s\right] D^2 \overline{w}(s) diag\left[s\right]\right)\right]$$

Proof. The proof is identical to the one in [58]. One has to use the fact that w is a viscosity solution of equation (3.4). The important assumption is that G is continuous, and that $\tilde{f}$ is continuous on its domain. In particular on domains such that $G > \varepsilon$. This is given by lemma 3.1. Given this lemma, the proof only relies on general PDE arguments. The only part that may become different is to prove that $\underline{v} \geq g$, because f may take infinite values. This is given in the following lemma. □

Lemma 3.2. *Suppose that g satisfy a linear growth condition and is lower semicontinuous. Then:*

$$\underline{w}(s) \geq g(s), \forall s \in \mathbb{R}_+^n$$

Proof. Take some arbitrary sequence $(t_m, s_m) \to (T, s)$ with $t_m < T$. By assumption 2.3 there exists a bounded continuous control $\sigma(t, S_t^{t_m,s_m}) \in \Sigma$ such that there exists a constant C verifying $f(t, S_t^m, \sigma(t, S_t^{t_m,s_m})) \leq C$ a.s. for every m. By definition of the value function we have:

$$\begin{aligned}
w(t_m, s_m) &\geq \mathbb{E}\left[\int_{t_m}^T -f(u, S_u^{t_m,s_m}, \sigma(t, S_t^{t_m,s_m}))du + g(S_T^{t_m,s_m})\right] \\
&\geq \mathbb{E}\left[\int_{t_m}^T -C + g(S_T^{t_m,s_m})\right] \\
&\geq -(T - t_m)C + \mathbb{E}\left[g(S_T^{t_m,s_m})\right]
\end{aligned}$$

Then, as σ is bounded we can use the dominated convergence theorem, and the linear growth of g to prove that:

$$\liminf_{m \to \infty} w(t_m, s_m) \geq \mathbb{E}\left[\liminf_{m \to \infty} g(S_T^{t_m,s_m})\right]$$

as g is lower-semicontinuous, and by the continuity of the flow $S_T^{t,s}$ in (t,s). □

The rest of the proofs are very classical and can be found in [58].

4 Equivalence with the dual formulation

In this section, we prove that the original value function in (2.2) is equal to the value function of the dual problem (2.3). This is achieved in two steps. In the first step, we show the inequality $v \geq w$. The second step : $v \leq w$ is more tedious to derive and involves PDE and hedging arguments. Let us begin by the straightforward part.

4.1 The inequality $v \geq w$

Proposition 4.4. *The super-replication price defined by (2.2) is larger than the value function of the dual problem (2.3). In other words:*

$$v \geq w \;\; on \;\; [0,T) \times \mathbb{R}^n$$

Proof. By definition of v, for any $(t,s) \in [0,T) \times \mathbb{R}^n$, and any $\varepsilon > 0$, there exist a hedging portfolio $\pi \in \mathcal{A}$ such that:

$$v(t,s) + \int_t^T \pi_u dS_u^{t,s,\sigma} \geq g(S_T^{t,s,\sigma}) - \int_t^T f(u, S_u^{t,s,\sigma}, \sigma_u^2) du,$$

for any process $\sigma \in \Sigma$. By definition of Σ and $\mathcal{A}$ one has that $\int_t^T \pi_u dS_u^\sigma$ is almost surely bounded from below. As is its a local martingale, it is therefore is a supermartingale, one can find that:

$$v(t,s) + \varepsilon \geq \mathbb{E}\left[g(S_T^{t,s,\sigma}) - \int_t^T f(u, S_u^{t,s,\sigma}, \sigma_u^2) du \right]$$

for any $\sigma \in \Sigma$. Hence $v(t,s) + \varepsilon \geq w(t,s)$ for any $\varepsilon > 0$. Letting $\varepsilon \to 0$ gives the required result. □

4.2 The inequality $v \leq w$

Now, let us show the reverse inequality. To prove it, we will use a hedging argument based on the Itô formula. In order to use this formula we need a $C^{1,2}$ function. As w is not supposed to be regular enough, we will need to use an approximation which is described as follows:

Definition 4.1. *For any $\varepsilon > 0$, let f_ε defined as:*

$$\begin{aligned} f_\varepsilon : [-\varepsilon, T] \times \mathbb{R}^d \times \mathcal{S}_d &\to \mathbb{R} \\ (t,s,\sigma^2) &\to \inf_{(t',s') \in B(t,s,\varepsilon)} \left\{ f(t',s',\sigma^2) \right\} \end{aligned}$$

where:

$$\begin{aligned} B(t,s,\varepsilon) = \Big\{ & (t',s') \in [0,T] \times \mathbb{R}^d \;\; s.t. \;\; t \leq t' \leq t + \varepsilon \\ & and \;\; s - \varepsilon \leq s' \leq s \;\; component \;\; wise \Big\} \end{aligned}$$

Let us consider the approximation of w, called w_ε, defined as the solution of the stochastic control problem:

$$w_\varepsilon(t,s) = \sup_{\sigma \in \Sigma} \mathbb{E}\left(g(S_T^{t,s,\sigma}) - \int_t^T f_\varepsilon(t', S_{t'}^{t,s,\sigma}, \sigma^2) dt' \right)$$

We see that the only difference between w and w_ε is that we replaced f by f_ε. As, by definition, $f_\varepsilon \leq f$, we obtain trivially that $w_\varepsilon \geq w$. In the following we will derive an upper bound on w_ε in order to show it is indeed a good approximation of w. But to derive that other bound, we have to use assumption 2.2. This is the most important reason why we had to introduce it. Now we can introduce another approximation of u which is regular:

Definition 4.2. *Let $\delta(t,s) \in C^\infty\left(\mathbb{R}^{d+1} \to \mathbb{R}\right)$ be a positive function such that $\int_{\mathbb{R}^{d+1}} \delta = 1$, and $\delta(x) = 0$ if $x \notin [0,1]^{d+1}$. And let the approximating function. $w^\varepsilon : [0,T) \times [0,+\infty)^n \to \mathbb{R}$ be defined as*

$$w^\varepsilon(t,s) = \int_{C(s,\varepsilon)} \int_{t-\varepsilon}^{t} w_\varepsilon(t',s') \frac{1}{\varepsilon^{n+1}} \delta(\frac{t-t'}{\varepsilon}, \frac{s-s'}{\varepsilon}) dt' ds' \tag{4.1}$$
$$\text{for all } (t,s) \in [0,T) \times [0,+\infty)^n$$

where $C(s,\varepsilon)$ is the set of points s' such that $s \leq s' \leq s+\varepsilon$.

Here, we prove the convergence of w^ε to w when $\varepsilon \to 0$:

Lemma 4.3. *For each $(t,s) \in [0,T) \times \mathbb{R}^d$ one has:*

$$\lim_{\varepsilon \to 0} w^\varepsilon(t,s) = w(t,s)$$

Proof. Indeed, with assumption 2.2, there exists a positive function h s.t. $\lim_{\varepsilon \to 0} h(\varepsilon) = 0$ and for any $(t',s') \in [0,T) \times \mathbb{R}^d$ and any process $\sigma \in \Sigma$:

$$f_\varepsilon(t',s',\sigma^2) \leq f(t',s',\sigma^2) \leq f_\varepsilon(t',s',\sigma^2) + h(\varepsilon)$$

Plugging this inequality in the definition of w and w_ε we find that, for any (t',s'):

$$w(t',s') \leq w_\varepsilon(t',s') \leq w(t',s') + h(\varepsilon) \tag{4.2}$$

Therefore, integrating (4.2) and using the definition of w^ε we get for any $(t,s,\varepsilon) \in [0,T) \times \mathbb{R}_+^d \times \mathbb{R}_+^*$:

$$\int_{C(s,\varepsilon)} \int_{t-\varepsilon}^{t} w(t',s') \frac{1}{\varepsilon^{n+1}} \delta(\frac{t-t'}{\varepsilon}, \frac{s-s'}{\varepsilon}) dt' ds' \leq w^\varepsilon(t,s) \leq$$
$$\int_{C(s,\varepsilon)} \int_{t-\varepsilon}^{t} w(t',s') \frac{1}{\varepsilon^{n+1}} \delta(\frac{t-t'}{\varepsilon}, \frac{s-s'}{\varepsilon}) dt' ds' + h(\varepsilon)$$

As w is continuous, (see proposition 5.6 below), the r.h.s and the l.h.s converge to $w(t,s)$ when $\varepsilon \to 0$, and the proof is completed. □

Now we restate a very classical result:

Lemma 4.4. *For any $\varepsilon > 0$, $w^\varepsilon \in C^\infty([0,T) \times [0,+\infty))$*

We can prove the following lemma which helps us to conclude that w^ε is actually a super-replication price:

Lemma 4.5. *For any $\varepsilon > 0$, w^ε is a supersolution of (3.4) with terminal condition $w^\varepsilon(T,.)$.*

Proof. We proceed as in [9], using the fact that the operator (3.4) is concave (see lemma 5.8 below). The steps of the proof are the following: As w_ε is continuous, one can approximate the integral with Riemann sums. These Riemann sum are supersolutions of (3.4) by lemma 5.8 . Then, as the Riemann sums converge uniformly to w^ε on any bounded domain, one can use the stability result for viscosity solutions (see [27] lemma 6.1), which completes the proof. □

This lemma provides a bound on the terminal condition $w^\varepsilon(T,.)$

Lemma 4.6. *There exists $K > 0$ such that $w^\varepsilon(T,s) + K\left(1 + |s|_1\right)\sqrt{\varepsilon} \geq g(s)$ for all $\varepsilon < 1$ and $s \in \mathbb{R}^n_+$*

Proof. By assumption 2.3, there exists a bounded function:

$$\begin{aligned} \hat{\sigma} : [0,T] \times \mathbb{R}^d &\to \mathcal{M}^d(\mathbb{R}) \\ (t,s) &\to \hat{\sigma}(t,s) \end{aligned}$$

And there exists K_1 s.t. $f(u,s,\hat{\sigma}^2(u,s)) \leq K_1$ and $|\hat{\sigma}(.,.)| \leq K_2$. By the arguments of lemma 3.2 we find that, for any $t \geq T - \varepsilon$:

$$\begin{aligned} w(t,s) &\geq E\left[g(S_T^{t,s,\hat{\sigma}}) - \int_t^T f(u,S_u,\hat{\sigma}^2(u,S_u))|\mathcal{F}_t\right] \\ &\geq E\left[g(S_T^{t,s,\hat{\sigma}})|\mathcal{F}_t\right] + K_1\varepsilon \end{aligned} \tag{4.3}$$

Using the Lipschitz condition on g, and the boundedness of $\hat{\sigma}$ we find that there exists K_3, depending of the Lipschitz coefficient of g, s.t. $E\left[g(S_T^{t,s',\hat{\sigma}})|\mathcal{F}_t\right] + K_2K_3\sqrt{\varepsilon} \geq g(s)$ for all $t > T - \varepsilon$ and $s > s' > s - \varepsilon$. Hence, combining this with inequality (4.3) we obtain that there exists K such that $w^\varepsilon(T,s) \geq g(s) - K\sqrt{\varepsilon}$ for all $\varepsilon < 1$. □

At last one can show that for each ε, the function $w^\varepsilon + K\left(1 + |s|_1\right)\sqrt{\varepsilon} \geq g(s)$ is greater that the super-replication price.

Proposition 4.5. *For each $(t,s) \in [0,T] \times \mathbb{R}^d$, $w(t,s) \geq v(t,s)$, that is, the auxiliary value function is greater that the super-replication price v. Together with proposition 4.4, one has $w = v$*

Proof. Let $(t',s') \in [0,T) \times \mathbb{R}^d$. Using the previous lemma, there exists K such that for any $1 > \varepsilon > 0$, then

$$w^\varepsilon(T,.) + K\left(1 + |.|_1\right)\sqrt{\varepsilon} \geq g(.) \tag{4.4}$$

Now, consider the function

$$\tilde{w}^\varepsilon(t,s) = w^\varepsilon(t,s) + K\left(1 + |s|_1\right)\sqrt{\varepsilon}$$

As, by lemma (4.4) $\tilde{w}^\varepsilon$ is regular enough, then for any a.s. bounded process σ_t we can use Itô's formula:

$$d\tilde{w}^\varepsilon(t,S_t) = D\tilde{w}^\varepsilon dS_t + \frac{\partial \tilde{w}^\varepsilon}{\partial t}dt + \frac{1}{2}tr(\mathrm{diag}[S_t]\sigma_t D^2\tilde{w}^\varepsilon\sigma_t\mathrm{diag}[S_t])dt.$$

Now, we introduce the following self-financed strategy:

$$X_T^{t',s'} = \int_{t'}^{T} -D\tilde{w}^{\varepsilon}(t, S_t^{t',s',\sigma_t})dS_t^{t',s',\sigma_t}.$$

Then, considering the initial wealth $\tilde{w}^{\varepsilon}(t', s')$, we obtain:

$$\begin{aligned} \tilde{w}^{\varepsilon}(t', s') + X_T^{t',s'} &= \tilde{w}^{\varepsilon}(T, S_T^{t',s',\sigma_T}) \\ &+ \int_{t'}^{T} \left(\frac{\partial w}{\partial t} + \frac{1}{2} tr(\mathrm{diag}[S_t^{\sigma_t}]\sigma_t D^2\tilde{w}^{\varepsilon}\sigma_t \mathrm{diag}[S_t^{\sigma_t}]) \right) dt. \end{aligned} \tag{4.5}$$

As $\tilde{w}^{\varepsilon}$ is a supersolution of 3.4, we get that:

$$-\frac{\partial \tilde{w}^{\varepsilon}}{\partial t} - (\mathrm{diag}[s](D^2\tilde{w}^{\varepsilon})\mathrm{diag}[s]) - f(t, s, \sigma^2) \geq 0.$$

Plugging this inequality into (4.5) gives:

$$\tilde{w}^{\varepsilon}(t', s') + X_T^{t',s'} \geq \tilde{w}^{\varepsilon}(T, S_T^{t',s',\sigma_T}) - \int_{t'}^{T} f(t, S_t^{t',s',\sigma_T}, \sigma^2)dt,$$

and so by (4.4), one finally gets:

$$\tilde{w}^{\varepsilon}(t', s') + X_T^{t',s'} \geq g(S_T^{t',s',\sigma_T}) - \int_{t'}^{T} f(t, S_t^{t',s',\sigma_T}, \sigma^2)dt.$$

This proves that $\tilde{w}^{\varepsilon}(t', s') \geq v(t', s')$. Now, letting $\varepsilon \to 0$, by definition of $\tilde{w}^{\varepsilon}$ and lemma 4.3 we obtain:

$$w(t', s') \geq v(t', s').$$

□

5 Comparison principle

In order to fully characterize the value function $v = w$ through PDE (3.4), we need an uniqueness result for that PDE. In this section we prove a comparison theorem for equation (3.4). It leads to the following uniqueness result:

Theorem 5.1. *Assume 2.2,2.3, 2.4 and 2.5. Then there exists at most one viscosity solution w of (3.4) with terminal condition $\hat{g}$ satisfiying:*

$$\hat{g} - C < w < \hat{g} + C,$$

for some constant C. Moreover, this function is continuous if it exists.

This theorem is proved at the end of the section.

5.1 Preliminaries

Regularity of the operator

Lemma 5.7. *For any $\varepsilon > 0$ there exist a constant K_ε, s.t. for any $(B, C) \in (\mathcal{S}_d)^2$, satisfying $B \geq C$ or $C \geq B$:*

$$\left|F(t,s,p,B) - F(t',s',p',C)\right| \leq \varepsilon + K_\varepsilon |sBs - s'Cs'| + |p-p'| + \beta(|s-s'| + |t-t'|), \quad (5.1)$$

where $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ is a function that does not depend on ε, B, and C, and $\beta(x) \to 0$ as $x \to 0$.

Proof. Let us decompose the inequality into several parts:

$$\begin{aligned} \left|F(t,s,p,B) - F(t',s',p,C)\right| \leq & \left|F(t,s,p,B) - F(t,s,p,D)\right| \\ & + \left|F(t,s,p,D) - F(t',s',p,C)\right|, \end{aligned}$$

where D is such that $\mathrm{diag}[s] D \mathrm{diag}[s] = \mathrm{diag}[s'] C \mathrm{diag}[s'] = \Gamma$. First, one can easely show that:

$$\left|F(t,s,p,D) - F(t',s',p,C)\right| \leq \beta(|s-s'| + |t-t'|). \quad (5.2)$$

Indeed, using assumption 2.2, we obtain:

$$|f(t,s,\sigma^2) - f(t',s',\sigma^2)| \leq \beta(|s-s'| + |t-t'|),$$

with $\beta = h^{-1}$ is the pseudo-inverse of h defined in assumption 2.2. Hence, by definition of $\tilde{f}$, if $\tilde{f}(t,s,\Gamma) < +\infty$:

$$\begin{aligned} \tilde{f}(t,s,\Gamma) &= \sup_{\sigma^2} \left\{ < \Gamma, \sigma^2 > - f(t,s,\sigma^2) \right\} \\ &\leq \sup_{\sigma^2} \left\{ < \Gamma, \sigma^2 > - f(t',s',\sigma^2) + \beta(|s-s'| + |t-t'|) \right\} \\ &\leq \tilde{f}(t,s,\Gamma) + \beta(|s-s'| + |t-t'|). \end{aligned}$$

Hence, if $\tilde{f}(t,s,\Gamma) < +\infty$, then:

$$|\tilde{f}(t,s,\Gamma) - \tilde{f}(t',s',\Gamma)| \leq \beta(|s-s'| + |t-t'|).$$

We conclude that (5.2) holds, reminding the definition of F, and noting that, due to lemma 3.1, G does only depend on:

$$\mathrm{diag}[s] D \mathrm{diag}[s] = \mathrm{diag}[s'] C \mathrm{diag}[s'] = \Gamma.$$

Now let us focus on the second inequality:

$$|F(t,s,p,B) - F(t,s,p,D)| \leq \varepsilon + K_\varepsilon |\mathrm{diag}[s'] D \mathrm{diag}[s'] - \mathrm{diag}[s'] C \mathrm{diag}[s']|.$$

With a symmetry argument, one has to prove actually the inequality:

$$F(t,s,p,B) - F(t,s,p,D) \geq -\varepsilon - K_\varepsilon |\mathrm{diag}[s'] D \mathrm{diag}[s'] - \mathrm{diag}[s'] C \mathrm{diag}[s']|.$$

By definition of F, we get:

$$\begin{aligned} F(t,s,p,B) - F(t,s,p,D) = &\sup_{A\geq 0}\left\{\min\left\{-\tilde{f}(t,s,\Gamma_1 - A), 1_{A=0}G(B) - tr(A)\right\}\right\} \\ &- \sup_{A\geq 0}\left\{\min\left\{-\tilde{f}(t,s,\Gamma_2 - A); 1_{A=0}G(D) - tr(A)\right\}\right\} \end{aligned}$$

Where we denoted $\Gamma_1 = \text{diag}[s]B\text{diag}[s]$ and $\Gamma_2 = \text{diag}[s]D\text{diag}[s]$. These matrices are symmetric, and so there exists $\Gamma_+ \geq 0$ and $\Gamma_- \leq 0$ s.t.:

$$\Gamma_2 = \Gamma_1 + \Gamma_+ + \Gamma_-$$

We have $\Gamma_2 - \Gamma_1 \geq \Gamma_-$. As F is a parabolic operator, one gets:

$$\begin{aligned} F(t,s,p,B) - F(t,s,p,D) \geq &\sup_{A\geq 0}\left\{\min\left\{-\tilde{f}(\Gamma_1 - A); 1_{A=0}G(\Gamma_1) - tr(A)\right\}\right\} \\ &- \sup_{A\geq 0}\left\{\min\left\{-\tilde{f}(\Gamma_1 + \Gamma_- - \hat{A}), 1_{A=0}G(\Gamma_1 + \Gamma_-) - tr(A)\right\}\right\}. \end{aligned}$$

Then by definition of the supremum, for any $\eta > 0$, there exists $\hat{A}$ such that:

$$\begin{aligned} F(t,s,p,B) - F(t,s,p,D) \geq &\sup_{A\geq 0}\left\{\min\left\{-\tilde{f}(\Gamma_1 - A); 1_{A=0}G(\Gamma_1) - tr(A)\right\}\right\} \\ &- \left\{\min\left\{-\tilde{f}(\Gamma_1 + \Gamma_- - \hat{A}), 1_{\hat{A}=0}G(\Gamma_1 + \Gamma_-) - tr(\hat{A})\right\}\right\} - \eta. \end{aligned}$$

Finally using the inequality $min(a,b) - min(c,d) \geq min(a-c, b-d)$ one gets:

$$\begin{aligned} F(t,s,p,B) - F(t,s,p,D) \geq &\sup_{A\geq 0}\left\{\min\left\{\tilde{f}(\Gamma_1 + \Gamma_- - \hat{A}) - \tilde{f}(\Gamma_1 - A),\right.\right. \\ &\left.\left. 1_{A=0}G(\Gamma_1) - 1_{\hat{A}=0}G(\Gamma_1 + \Gamma_-) + tr(\hat{A} - A)\right\}\right\} - \eta \end{aligned}$$

Now we divide the proof in two cases, depending on whether $G(\Gamma_1) \geq \varepsilon$ or $G(\Gamma_1) \leq \varepsilon$:

- *Case* $G(\Gamma_1) \leq \varepsilon$:
  Taking $A = -\Gamma_- + \hat{A}$ leads to:

$$\begin{aligned} F(t,s,p,B) - F(t,s,p,D) \geq \min &\left\{\tilde{f}(\Gamma_1 + \Gamma_- - \hat{A}) - \tilde{f}(\Gamma_1 - A),\right. \\ &\left. 1_{A=0}G(\Gamma_1) - 1_{\hat{A}=0}G(\Gamma_1 + \Gamma_-) + tr(\hat{A} - A)\right\} - \eta \end{aligned}$$

  and we get:

$$F(t,s,p,B) - F(t,s,p,D) \geq \min\left\{0, -1_{\hat{A}=0}G(\Gamma_1 + \Gamma_-) + tr(\Gamma_-)\right\} - \eta.$$

  Using the K-Lipschitz continuity of G given by lemma 3.1, we have for all $\eta > 0$:

$$\begin{aligned} F(t,s,p,B) - F(t,s,p,D) &\geq \min\left\{0, -1_{\hat{A}=0}(G(\Gamma_1 + \Gamma_-) - G(\Gamma_1))\right. \\ &\quad\left. - 1_{\hat{A}=0}G(\Gamma_1) + tr(\Gamma_-)\right\} - \eta \\ F(t,s,p,B) - F(t,s,p,D) &\geq -(K+1)|\Gamma_-| - \varepsilon - \eta \\ F(t,s,p,B) - F(t,s,p,D) &\geq -(K+1)|\text{diag}[s']D\text{diag}[s'] \\ &\quad - \text{diag}[s']B\text{diag}[s']| - \varepsilon - \eta \end{aligned}$$

- *Case* $G(\Gamma_1) \geq \varepsilon$:
  Taking $A = \hat{A}$ and using assumption 2.4 gives:

$$F(t,s,p,B) - F(t,s,p,D) \geq \min\Big\{\tilde{f}(\Gamma_1 + \Gamma_- - \hat{A}) - \tilde{f}(\Gamma_1 - \hat{A}),$$
$$1_{A=0}(G(\Gamma_1) - G(\Gamma_1 + \Gamma_-))\} - \eta$$
$$F(t,s,p,B) - F(t,s,p,D) \geq -max(K_\varepsilon, K)|\Gamma_-| - \eta$$
$$F(t,s,p,B) - F(t,s,p,D) \geq -max(K_\varepsilon, K)|\mathrm{diag}[s']D\mathrm{diag}[s'] - \mathrm{diag}[s']B\mathrm{diag}[s']| - \eta.$$

Taking the limit when $\eta \to 0$ gives the required result. □

Strict supersolution

First, we must prove the concavity of the operator with respect to the test functions:

Lemma 5.8. *For any $(t,s) \in [0,T) \times R_+^n$, let φ_1 and φ_2 two C^2 test functions, such that:*

$$F(t,s,D\varphi_1,D^2\varphi_1) \geq 0 \textit{ and } F(t,s,D\varphi_2,D^2\varphi_2) \geq 0.$$

Then for any $\lambda \in [0,1]$, denoting $\varphi = \lambda\varphi_1 + (1-\lambda)\varphi_2$, we have:

$$F(t,s,D\varphi,D^2\varphi) \geq \lambda F(t,s,D\varphi_1,D^2\varphi_1) + (1-\lambda)F(t,s,D\varphi_2,D^2\varphi_2). \tag{5.3}$$

Proof. Let us fix (t,s) and λ for the rest of the proof. First of all, it is obvious that $\tilde{f}(t,s,\mathrm{diag}\,[s]\,D^2\varphi\mathrm{diag}\,[s])$ is convex in its last argument, as it is the convex conjugate of f. Then, for any $n \times n$ symmetric matrices A_1 and A_2:

$$\tilde{f}(t,s,D^2\varphi + \lambda A_1 + (1-\lambda)A_2) \leq \lambda\tilde{f}(t,s,D^2\varphi_1 + A_1) + (1-\lambda)\tilde{f}(t,s,D^2\varphi_2 + A_2). \tag{5.4}$$

Lemma 3.1 proves that $G(t,s,D^2\varphi)$ is concave w.r.t its third variable. First, we will prove that, when the operator is positive, then the supremum in its definition is attained for $A = 0$. By definition of F, there exists a sequence $A_n \geq 0$ s.t.:

$$F(t,s,D\varphi_1,D^2\varphi_1) = \lim\Bigg(\min\Bigg\{\frac{\partial\varphi_1}{\partial t} - \tilde{f}\left(\mathrm{diag}\,[s]\,D^2\varphi_1\mathrm{diag}\,[s] - A_n\right),$$
$$1_{A_n=0}G(\mathrm{diag}\,[s]\,D^2\varphi_1\mathrm{diag}\,[s]) - tr(A_n)\Big\}\Big) \geq 0.$$

Looking at the second term of the minimum, we notice that $A_n \to 0$. As $\tilde{f}$ is lower semicontinuous w.r.t. its third variable, one gets:

$$\frac{\partial\varphi_1}{\partial t} - \tilde{f}\left(\mathrm{diag}\,[s]\,D^2\varphi_1\mathrm{diag}\,[s]\right) \geq F(t,s,D\varphi_1,D^2\varphi_1).$$

Moreover, it is obvious that:

$$G(\mathrm{diag}\,[s]\,D^2\varphi_1\mathrm{diag}\,[s]) \geq F(t,s,D\varphi_1,D^2\varphi_1).$$

Hence the supremum is indeed attained for $A = 0$. The same arguments can be applied for φ_2. As $\tilde{f}(t, s, .)$ and $G(.)$ are concave functions, so is their minimum. Finally we get:

$$\begin{aligned} F(t, s, D\varphi, D^2\varphi) &\geq \min\left\{\frac{\partial \varphi}{\partial t} - \tilde{f}\left(\text{diag}\,[s]\, D^2\varphi \text{diag}\,[s]\right), G(\text{diag}\,[s]\, D^2\varphi \text{diag}\,[s])\right\} \\ &\geq \lambda F(t, s, D\varphi_1, D^2\varphi_1) + (1-\lambda) F(t, s, D\varphi_2, D^2\varphi_2), \end{aligned}$$

and this concludes the proof. □

Let us point out a η-strict supersolution of the equation:

Lemma 5.9. *Let assumption 2.4 hold. Then, for any constant c^* the function:*

$$w^1(t, s) = (T - t) + c^*$$

is a η^1-strict supersolution of (3.4) for some $\eta^1 > 0$.

Proof. It follows from assumption 2.4 which ensures existence of $\eta^1 > 0$ such that $G(t, 0) > \eta^1$ for all $t \in [0, T]$. Hence, taking $A = 0$ in (3.4) gives:

$$F(t, s, w_t^1(t, s), D^2 w^1(t, s)) \geq min(1, \eta^1) \forall (t, s) \in [0, T] \times R_+^d.$$

□

One can also easily build η-strict supersolutions of (3.4), as in [23].

Lemma 5.10. *Let w^0 be a lower semicontinuous viscosity supersolution of the equation:*

$$F(t, s, w_t^0(t, s), D^2 w^0(t, s)) = 0, \tag{5.5}$$

and w^1 be a lower semicontinuous η-strict supersolution of the equation (5.5) for some $\eta \geq 0$. Then, for all $\mu \in (0, 1)$, the function $w^\mu := (1-\mu)w^0 + \mu w^1$ is a $\mu\eta$-strict supersolution of equation (5.5).

Proof. The proof is exactly the same as in [23], which is based on the concavity of F with respect to the solution. □

5.2 Comparison result

Proposition 5.6. *Let assumptions 2.3, 2.2 and 2.4 hold. Suppose u is an upper semicontinuous subsolution of (3.4) and w a lower semicontinuous η-strict viscosity supersolution of (3.4) for some $\eta > 0$. Assume that there exists a constant C such that:*

$$u(t, s) \leq \hat{g}(s) + C \text{ and } w(t, s) \geq \hat{g}(s) - C \text{ for all } (t, s) \in [0, T) \times (0, +\infty)^d \tag{5.6}$$

Then, $u(T, .) \leq \hat{g}(.) \leq w(T, .)$ implies $u(t, s) \leq w(t, s)$ on $[0, T) \times (0, +\infty)^d$.

Proof. We adapt the comparison result of [23]. The difference is that, here, one does not have the Lipschitz condition on F, but a weaker condition (5.1) given by lemma 5.7. As we show here, it does not interfere in the proof.
For $\varepsilon, \alpha > 0$, consider the upper semicontinuous function:

$$\Phi^{(\varepsilon,\alpha)}(t,t',s,s') = u(t,s) - w(t',s') - \varepsilon(l(s) + l(s')) - \frac{1}{2}\alpha\left((t-t')^2 + (s-s')^2\right)$$

where

$$l(s) = \sum_{j=1}^{d} [s_j - \log s_j].$$

Let us consider the function:

$$\Phi^{\varepsilon}(t,s) = \Phi^{(\varepsilon,\alpha)}(t,t,s,s).$$

Then, with condition (5.6), we see that ϕ^{ε} is bounded from above and tends to $-\infty$ on the boundary of the domain, hence there exist $t_\varepsilon, s_\varepsilon$ such that:

$$\max_{[0,T]\times[0,\infty)^d} \phi^{\varepsilon}(t,s) = \phi^{\varepsilon}(t,s).$$

Now, let us study the case when:

$$t_{\varepsilon_k} = T \text{ for some sequence } (\varepsilon_k)_{k\geq 1} \text{ with } \varepsilon_k > 0 \text{ and } \varepsilon_k \to 0. \tag{5.7}$$

Exactly as in [23] one can arrive at:

$$\begin{aligned} u(t,s) - w(t,s) &= \Phi^{\varepsilon_k}(t,s) + 2\varepsilon_k l(s) \\ &\leq \Phi^{\varepsilon_k}(t, s_{\varepsilon_k}) + 2\varepsilon_k l(s) \\ &\leq u(T, s_{\varepsilon_k}) - w(T, s_{\varepsilon_k}) - 2\varepsilon_k l(s_{\varepsilon_k}) + 2\varepsilon_k l(s) \\ &\leq u(T, s_{\varepsilon_k}) - w(T, s_{\varepsilon_k}) + 2\varepsilon_k l(s) \\ &\leq 2\varepsilon_k l(s) \text{ for all } (t,s) \in [0,T) \times (0,+\infty)^d. \end{aligned}$$

Taking $k \to \infty$ gives the required result.
Now, in order to prove (5.7), we assume on the contrary that there exists a constant $\overline{\varepsilon} > 0$ such that

$$t_\varepsilon < T \text{ for all } 0 < \varepsilon \leq \overline{\varepsilon}$$

and we work toward a contradiction. Using the same arguments as in [23], one can prove that there exists a sequence $\alpha_k \to \infty$ and a sequence $(t_k, t'_k, s_k, s'_k) \to (t_\varepsilon, t'_\varepsilon, s_\varepsilon, s'_\varepsilon)$ such that

$$\alpha_k\left((t_k - t'_k)^2 + (s_k - s'_k)^2\right) \to 0 \text{ as } k \to \infty.$$

Note that for ε sufficiently small and α_k sufficiently large, $t_k < T$ and $t'_k < T$. Based on theorem 3.2 in [27], the arguments in [23] lead, for sufficiently large α_k, to existence of two symmetric matrices $A_k, A'_k \in \mathcal{S}^d$ such that:

$$-3\alpha_k \begin{pmatrix} I_d & 0 \\ 0 & I_d \end{pmatrix} \leq \begin{pmatrix} A_k & 0 \\ 0 & -A'_k \end{pmatrix} \leq 3\alpha_k \begin{pmatrix} I_d & -I_d \\ -I_d & I_d \end{pmatrix} \tag{5.8}$$

and

$$F\left(s_k, p_k + (t_k - t_\varepsilon), A_k + \varepsilon D^2 l(s_k) + Q(s_k - s_\varepsilon)\right) \leq 0$$
$$F\left(s_k, p_k - (t_k - t_\varepsilon), A'_k - \varepsilon D^2 l(s'_k) - Q(s'_k - s_\varepsilon)\right) \geq \eta$$

where

$$p_k := \alpha_k(t_k - t'_k) \text{ and } Q(z) := 2z \otimes z + |z|^2 Id.$$

First, we calculate explicitly the norms of the quantities defined above:

$$D^2 l(s_k) = \mathrm{diag}[\frac{1}{s_k^2}]$$

and

$$|Q(s_k - s_\varepsilon)| \leq 2|(s_k - s_\varepsilon).(s_k - s_\varepsilon)'| + |(s_k - s_\varepsilon)'.(s_k - s_\varepsilon)|.$$

Then, we use lemma 5.7 with $\frac{\eta}{6}$ to conclude that:

$$\begin{aligned} F(s_k, p_k, A_k) &\leq F\left(s_k, p_k + (t_k - t_\varepsilon), A_k + \varepsilon D^2 l(s_k) + Q(s_k - s_\varepsilon)\right) \\ &\quad + \frac{\eta}{6} + |t_k - t_\varepsilon| + K_{\frac{\eta}{6}} |\mathrm{diag}[s_k][\varepsilon D^2 l(s_k) + Q(s_k - s_\varepsilon)]\mathrm{diag}[s_k]| \\ &\leq \frac{\eta}{6} + |t_k - t_\varepsilon| + K_{\frac{\eta}{6}} |\varepsilon + 3|s_k - s_\varepsilon|^2 |s_k|^2|. \end{aligned}$$

Using the same arguments, one can prove that:

$$F(s'_k, p_k, A'_k) \geq \eta - \frac{\eta}{6} - |t'_k - t_\varepsilon| - K_{\frac{\eta}{6}} |\varepsilon + 3|s'_k - s_\varepsilon|^2 |s'_k|^2|.$$

Combining these two inequalities, we get:

$$\begin{aligned} F(s'_k, p_k, A'_k) - F(s_k, p_k, A_k) &\geq \eta - \frac{\eta}{3} - |t'_k - t_\varepsilon| - |t_k - t_\varepsilon| \\ &\quad - 3K_{\frac{\eta}{6}} |\varepsilon + |s_k - s_\varepsilon|^2 |s_k|^2 + |s'_k - s_\varepsilon|^2 |s'_k|^2|. \end{aligned}$$

Therefore, letting $k \to \infty$:

$$\liminf_{k \to \infty} \left[F(s'_k, p_k, A'_k) - F(s_k, p_k, A_k)\right] \geq \frac{2\eta}{3} - 2K_{\frac{\eta}{6}} \varepsilon. \tag{5.9}$$

On the other hand, since A_k, A'_k satisfy (5.8), we get, by multiplying both sides by $(\mathrm{diag}[s_k]\mathrm{diag}[s'_k])$ and $\begin{pmatrix} \mathrm{diag}[s_k] \\ \mathrm{diag}[s'_k] \end{pmatrix}$:

$$\mathrm{diag}[s_k] A_k \mathrm{diag}[s_k] - \mathrm{diag}[s'_k] A'_k \mathrm{diag}[s'_k] \leq diag\left((s_k - s'_k)^2\right)$$

Finally, using lemma 5.7 with $\frac{\eta}{3}$, we get:

$$\begin{aligned} F(s'_k, p_k, A'_k) - F(s_k, p_k, A_k) &\leq \frac{\eta}{3} + K_{\frac{\eta}{3}} |\mathrm{diag}[s_k] A_k \mathrm{diag}[s_k] - \mathrm{diag}[s'_k] A'_k \mathrm{diag}[s'_k]| \\ &\quad + \beta(|s_k - s'_k|) \\ &\leq \frac{\eta}{3} + K_{\frac{\eta}{3}} \left(\alpha_k[|s_k - s'_k|^2]\right) + \beta(|s_k - s'_k|). \end{aligned} \tag{5.10}$$

Since $\alpha_k[|s_k - s'_k|^2] \to 0$ anf $|s_k - s'_k| \to 0$ as $k \to \infty$, we obtain a contradiction between (5.9) and (5.10) when $\varepsilon < \frac{\eta}{6K_{\frac{\eta}{6}}}$. Hence (5.7) has to hold, and comparison is proved. □

5.3 Proof of theorem 5.1

Let u and w be two viscosity solutions of (3.4). Then we know with lemma 5.10, that $w^\mu = w + \mu w^1$, where w^1 is defined in lemma 5.9, is a $\mu\eta$ strict supersolution of (3.4). Furthermore, with assumption 2.5 we get the existence of a constant C' such that $\hat{g} - C' \leq w^\mu$. Hence, using proposition 5.6 we get $w^\mu_* \geq u^*$ for all $\mu > 0$. Taking $\mu \to 0$ we get $w_* \geq u^*$. By a symmetric argument, we get $u_* \geq w^*$. Moreover, as we have $u^* \geq u_*$, we finally get $w_* \geq u^* \geq u_* \geq w^*$, and therefore $w = u$ and w is continuous.

Part III

Impulse control problem with delay

Chapter 4

Impulse control problem on finite horizon with execution delay

This is a joint work with Huyen Pham [1]

We consider impulse control problems in finite horizon for diffusions with decision lag and execution delay. The new feature is that our general framework deals with the important case when several consecutive orders may be decided before the effective execution of the first one. This is motivated by financial applications in the trading of illiquid assets such as hedge funds. We show that the value functions for such control problems satisfy a suitable version of dynamic programming principle in finite dimension, which takes into account the past dependence of state process through the pending orders. The corresponding Bellman partial differential equations (PDE) system is derived, and exhibit some peculiarities on the coupled equations, domains and boundary conditions. We prove a unique characterization of the value functions to this nonstandard PDE system by means of viscosity solutions. The uniqueness result and the boundary conditions are obtained by backward and forward iterations on the domains and the value functions.

Key words : Impulse control, execution delay, diffusion processes, dynamic programming, viscosity solutions, comparison principle.

[1]Laboratoire de Probabilités et Modèles Aléatoires, CNRS, UMR 7599, Université Paris 7 Diderot, and Institut Universitaire de France, pham@math.jussieu.fr

1 Introduction

In this paper, we consider a general impulse control problem in finite horizon of a diffusion process X, with intervention lag and execution delay. This means that we may intervene on the diffusion system at any times τ_i separated at least by some fixed positive lag h, by giving some impulse ξ_i based on the information at τ_i. However, the execution of the impulse decided at τ_i is carried out with delay mh, $m \geq 1$, i.e. it is implemented at time $\tau_i + mh$, moving the system from $X_{(\tau_i+mh)^-}$ to $\Gamma(X_{(\tau_i+mh)^-}, \xi_i)$. The objective is to maximize over impulse controls $(\tau_i, \xi_i)_i$ the expected total profit on finite horizon T, of the form

$$\mathbb{E}\Big[\int_0^T f(X_t)dt + g(X_T) + \sum_{\tau_i+mh\leq T} c(X_{(\tau_i+mh)^-}, \xi_i)\Big].$$

Such formulations appear naturally in decision-making problems in economics and finance. In many situations, firms or investors face regulatory delays (delivery lag), which may be significant, and thus need to be taken into account when management strategies are decided in an uncertain environment. Problems where firm's investment are subject to delivery lag can be found in the real options literature, for example in [4] and [2]. In financial market context, execution delay is related to liquidity risk (see e.g. [66]), and occurs with transaction, which requires heavy preparatory work as for hedge funds. Indeed, hedge funds frequently hold illiquid assets, and need some time to find a counterpart to buy or sell them. Furthermore, this notice period gives the hedge fund manager a reasonable investement horizon.

From a mathematical viewpoint, it is well-known that impulse control problems without delay, i.e. $m = 0$, lead to variational partial differential equations (PDE), see e.g. the books [14] and [55]. Impulse control problems in the presence of delay were studied in [60] for $m = 1$, that is when no more than one pending order is allowed at any time. In this case, it is shown that the delay problem may be transformed into a no-delay impulse control problem. The paper [11] also considers the case $m = 1$, but when the value of the impulse is chosen at the time of execution, and on infinite horizon, and these two conditions are crucial in the proposed probabilistic resolution. We mention also the works [4] and recently [55], which study impulse problems in infinite horizon with arbitrary number of pending orders, but under restrictive assumptions on the controlled state process, like (geometric) Lévy process for X and (multiplicative) additive intervention operator Γ. In this case, the problem is reduced to a finite-dimensional one where the value functions with pending orders are directly related to the value function without order.

The main contribution of this paper is to provide a theory of impulse control problems with delay on finite horizon in a fairly general diffusion framework that deals with the important case in applications when the number of pending orders is finite, but not restricted to one, i.e. $m \geq 1$. Our chief goal is to obtain a unique tractable PDE characterization of the value functions for such problems. As usual in stochastic control problems, the first step is the derivation of a dynamic programming principle (DPP). We show a suitable version of DPP, which takes into account the past dependence of the controlled diffusion via the finite number of pending orders. The corresponding Bellman PDE system reveals some nonstandard features both on the form of the differential operators and their domains, and on the boundary conditions. Following the modern approach to stochastic control, we prove that the value functions are viscosity solutions to this Bellman PDE system, and we also state comparison principles, which allows to obtain a unique PDE characterization. From this PDE representation, we will provide in the next chapter an easily implemented algorithm to compute the value functions, and so as byproducts the optimal impulse control. This algorithm involves forward and backward iterations on the value functions and on the domains, and appear actually as original arguments in the proofs for the boundary conditions and comparison principles.

The rest of the paper is organized as follows. In Section 2, we formulate the control problem and introduce the associated value functions. Section 3 deals with the dynamic programming principle in this general framework. We then state in Section 4 the unique PDE viscosity characterization for the value functions. In section 5, we describe how to derive the optimal control from the value functions. Finally, Section 6 is devoted to the proofs of results in this paper.

2 Problem formulation

2.1 The control problem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions, and $W = (W_t)_{t\geq 0}$ a standard n-dimensional Brownian motion.

An impulse control is a double sequence $\alpha = (\tau_i, \zeta_i)_{i\geq 1}$, where (τ_i) is an increasing sequence of $\mathbb{F}$-stopping times, and ξ_i are $\mathcal{F}_{\tau_i}$-measurable random variables valued in E. We require that $\tau_{i+1} - \tau_i \geq h$ a.s., where $h > 0$ is a fixed time lag between two decision times, and we assume that E, the set of impulse values, is a compact subset of $\mathbb{R}^q$. We denote by $\mathcal{A}$ this set of impulse controls.

In absence of impulse executions, the system valued in $\mathbb{R}^d$ evolves according to :

$$dX_s \quad = \quad b(X_s)ds + \sigma(X_s)dW_s, \tag{2.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma \to \mathbb{R}^{d\times n}$ are Borel functions on $\mathbb{R}^d$, satisfying usual Lipschitz conditions. The interventions are decided at times τ_i with impulse values ξ_i based on the information at these dates, however they are executed with delay at times $\tau_i + mh$, moving the system from $X_{(\tau_i+mh)^-}$ to $X_{(\tau_i+mh)} = \Gamma(X_{(\tau_i+mh)^-}, \xi_i)$. Here Γ is a mapping from $\mathbb{R}^d \times E$ into $\mathbb{R}^d$, and we assume that Γ is continuous, and satisfies the linear growth condition :

$$\sup_{(x,e)\in\mathbb{R}^d\times E} \frac{|\Gamma(x,e)|}{1+|x|} \quad < \quad \infty. \tag{2.2}$$

Given an impulse control $\alpha = (\tau_i, \xi_i)_{i\geq 1} \in \mathcal{A}$, and an initial condition $X_0 \in \mathbb{R}^d$, the controlled process X^α is then defined as the solution to the s.d.e. :

$$X_s \quad = \quad X_0 + \int_0^s b(X_u)du + \int_0^s \sigma(X_u)dW_u + \sum_{\tau_i+mh\leq s} \big(\Gamma(X_{(\tau_i+mh)^-}, \xi_i) - X_{(\tau_i+mh)^-}\big). \tag{2.3}$$

We now fix a finite horizon $T < \infty$, and in order to avoid trivialities, we assume $T - mh \geq 0$. Using standard arguments based on Burkholder-Davis-Gundy's inequality, Gronwall's lemma and (2.2), we easily check that

$$\mathbb{E}\big[\sup_{s\leq T}|X_s^\alpha|\big] \quad < \quad \infty. \tag{2.4}$$

Given an impulse control $\alpha = (\tau_i,\xi_i)_{i\geq 1} \in \mathcal{A}$, we consider the total profit at horizon T, defined by :

$$\Pi(\alpha) \quad = \quad \int_0^T f(X_s^\alpha)ds + g(X_T^\alpha) + \sum_{\tau_i+mh\leq T} c(X^\alpha_{(\tau_i+mh)^-},\xi_i),$$

and we assume that the running profit function f, the terminal profit function g, and the executed cost function c are continuous, and satisfy the linear growth condition :

$$\sup_{(x,e)\in\mathbb{R}^d\times E} \frac{|f(x)|+|g(x)|+|c(x,e)|}{1+|x|} \quad < \quad \infty. \tag{2.5}$$

This ensures with (2.4) that $\Pi(\alpha)$ is integrable, and we can define the control problem :

$$V_0 \quad = \quad \sup_{\alpha\in\mathcal{A}} \mathbb{E}\big[\Pi(\alpha)\big]. \tag{2.6}$$

We also impose the following assumption :

$$g(x) \quad \geq \quad g\big(\Gamma(x,e)\big) + c(x,e), \qquad \forall\ (x,e) \in \mathbb{R}^d \times E. \tag{2.7}$$

This condition economically means that a decision at time $T-mh$ induces a terminal profit, which is smaller than a no-decision at this time $T-mh$, and is thus suboptimal. Mathematically, we shall see later that the condition (2.7) is crucial for the continuity of the value function associated to our problem, see Remark 4.2 **3**. Finally, notice that any intervention decided after date $T-mh$ will not influence the system and so the total profit at horizon T, and therefore, we may require w.l.o.g. that any admissible impulse control α $= (\tau_i,\xi_i)_{i\geq 1} \in \mathcal{A}$ satisfies $\tau_i + mh \leq T$ for all i s.t. $\tau_i < \infty$.

Financial example

Consider a financial market consisting of a money market account yielding a constant interest rate r, and a risky asset (stock) of price process $(S_t)_t$ governed by :

$$dS_t \quad = \quad \beta(S_t)dt + \gamma(S_t)dW_t.$$

We denote by Y_t the number of shares in the stock, and by Z_t the amount of money (cash holdings) held by the investor at time t. We assume that the investor can only trade discretely, and her orders are executed with delay. This is modelled through an impulse control $\alpha = (\tau_i,\xi_i)_{i\geq 1} \in \mathcal{A}$, where τ_i are the decision times, and ξ_i are the numbers of stock purchased if $\xi_i \geq 0$ or selled if $\xi_i < 0$ decided at τ_i, but executed at times $\tau_i + mh$. The dynamics of Y is then given by

$$Y_t \quad = \quad Y_0 + \sum_{\tau_i+mh\leq t} \xi_i,$$

which means that discrete trading $\Delta Y_t := Y_t - Y_{t^-} = \xi_i$ occur at times $s = \tau_i + mh$, $i \geq 1$. In absence of trading, the cash holdings Z grows deterministically at rate r : $dZ_t = rZ_tdt$.

When a discrete trading ΔY_t occurs, this results in a variation of cash holdings by $\Delta Z_t := Z_t - Z_{t^-} = -(\Delta Y_t) S_t$, from the self-financing condition. In other words, the dynamics of Z is given by

$$Z_t \quad = \quad Z_0 + \int_0^t r Z_u du - \sum_{\tau_i + mh \le t} \xi_i . S_{\tau_i + mh}.$$

The wealth process is equal to $L(S_t, Y_t, Z_t) = Z_t + Y_t S_t$. This financial example corresponds to the general model (2.3) with $X = (S, Y, Z)$, $b = (\beta\ 0\ r)'$, $\sigma = (\gamma\ 0\ 0)$, and

$$\Gamma(s, y, z, e) \quad = \quad \begin{pmatrix} s \\ e \\ z - es \end{pmatrix}.$$

In this case, condition (2.7) is satisfied with an equality . Fix now some contingent claim characterized by its payoff at time T : $H(S_T)$ for some measurable function H. The two following hedging and valuation criteria are very popular in finance, and may be embedded in our general framework :

• *Shortfall risk hedging.* The investor is looking for a trading strategy that minimizes the shortfall risk of the $P\&L$ between her contingent claim and her terminal wealth,

$$\inf_{\alpha \in \mathcal{A}} \mathbb{E}\Big[\Big(H(S_T) - L(S_T, Y_T, Z_T)\Big)_+\Big].$$

• *Utility indifference price.* Given an utility function U for the investor, an initial capital z in cash, zero in stock, and $\kappa \ge 0$ units of contingent claims, define the expected utility under optimal trading

$$V_0(z, \kappa) \quad = \quad \sup_{\alpha \in \mathcal{A}} \mathbb{E}\Big[U\big(L(S_T, Y_T, Z_T) - \kappa H(S_T)\big)\Big].$$

The utility indifference ask price $\pi_a(\kappa, z)$ is the price at which the investor is indifferent (in the sense that her expected utility is unchanged under optimal trading) between paying nothing and not having the claim, and receiving $\pi_a(\kappa, z)$ now to deliver κ units of claim at time T. It is then defined as the solution to

$$V_0(z + \pi_a(\kappa, z), \kappa) \quad = \quad V_0(z, 0).$$

2.2 Value functions

In order to provide an analytic characterization of the control problem (2.6), we need as usual to extend the definition of this control problem to general initial conditions. However, in contrast with classical control problems without execution delay, the diffusion process solution to (2.3) is not Markovian. Actually, given an impulse control, we see that the state of the system is not only defined by its current state value at time t but also by the pending orders, that is the orders not yet executed, i.e. decided between time $t - mh$ and t. Notice that the number of pending orders is less or equal to m. Let us then introduce the following definitions and notations. For any $t \in [0, T]$, $k = 0, \ldots, m$, we denote by

$$\begin{aligned} P_t(k) \quad = \quad \Big\{ & p = (t_i, e_i)_{1 \le i \le k} \in ([0, T - mh] \times E)^k \ : t_i - t_{i-1} \ge h, \ \ i = 2, \ldots, k, \\ & t - mh < t_i \le t, \ i = 1, \ldots, k \Big\}, \end{aligned}$$

the set of k pending orders not yet executed before time t, with the convention that $P_t(0) = \emptyset$. For any $p = (t_i, e_i)_{1\leq i\leq k} \in P_t(k)$, $t \in [0,T]$, $k = 0, \ldots, m$, we denote

$$\mathcal{A}_{t,p} \quad = \quad \Big\{\alpha = (\tau_i, \xi_i)_{i\geq 1} \in \mathcal{A} \ : (\tau_i, \xi_i) = (t_i, e_i), \ i = 1, \ldots, k \ \text{ and } \ \tau_{k+1} \geq t\Big\},$$

the set of admissible impulse controls with pending orders p before time t.

For any $(t,x) \in [0,T] \times \mathbb{R}^d$, $p \in P_t(k)$, $k = 0, \ldots, m$, and $\alpha \in \mathcal{A}_{t,p}$, we denote by $X^{t,x,p,\alpha}$ the solution to (2.3) for $t \leq s \leq T$, with initial data $X_t = x$, and pending orders p, i.e.

$$X_s \quad = \quad x + \int_t^s b(X_u)du + \int_t^s \sigma(X_u)dW_u + \sum_{t<\tau_i+mh\leq s} \big(\Gamma(X_{(\tau_i+mh)^-}, \xi_i) - X_{(\tau_i+mh)^-}\big).$$

Using standard arguments based on Burkholder-Davis-Gundy's inequality, Gronwall's lemma and (2.2), we easily check that

$$\mathbb{E}\big[\sup_{t\leq s\leq T} |X_s^{t,x,p,\alpha}|^2\big] \quad \leq \quad C(1+|x|^2), \tag{2.8}$$

for some positive constant C depending only on b, σ, Γ and T. We then consider the following performance criterion :

$$J_k(t,x,p,\alpha) \quad = \quad \mathbb{E}\Big[\int_t^T f(X_s^{t,x,p,\alpha})ds + g(X_T^{t,x,p,\alpha}) + \sum_{t<\tau_i+mh\leq T} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-}, \xi_i)\Big],$$

for $(t,x) \in [0,T] \times \mathbb{R}^d$, $p \in P_t(k)$, $k = 0, \ldots, m$, $\alpha = (\tau_i, \xi_i)_i \in \mathcal{A}_{t,p}$, and the corresponding value functions :

$$v_k(t,x,p) \quad = \quad \sup_{\alpha\in\mathcal{A}_{t,p}} J_k(t,x,p,\alpha), \qquad k = 0, \ldots, m, \ (t,x,p) \in \mathcal{D}_k,$$

where $\mathcal{D}_k$ is the definition domain of v_k :

$$\mathcal{D}_k \quad = \quad \big\{(t,x,p) \ : (t,x) \in [0,T) \times \mathbb{R}^d, \ p \in P_t(k)\big\}.$$

For $k = 0$, $P_t(0) = \emptyset$, and we write by convention $v_0(t,x) = v_0(t,x,\emptyset)$, $\mathcal{D}_0 = [0,T) \times \mathbb{R}^d$ so that the original control problem in (2.6) is given by $V_0 = v_0(0, X_0)$. Note, however, that v_0 is defined on $[0,T] \times \mathbb{R}^d$. Notice from (2.5) and (2.8) that the functions v_k satisfy the linear growth condition on $\mathcal{D}_k$:

$$\sup_{(t,x,p)\in\mathcal{D}_k} \frac{|v_k(t,x,p)|}{1+|x|} \quad < \quad \infty, \qquad k = 0, \ldots, m. \tag{2.9}$$

3 Dynamic programming

In this section, we state the dynamic programming relation on the value functions of our control problem with delay execution. For any $t \in [0,T]$, $\alpha = (\tau_i, \xi_i)_{i\geq 1} \in \mathcal{A}$, we denote :

$$\iota(t,\alpha) \quad = \quad \inf\{i \geq 1 \ : \tau_i > t - mh\} - 1 \ \in \ \mathbb{N} \cup \{\infty\}, \tag{3.1}$$

$$k(t,\alpha) \quad = \quad \text{card}\big\{i \geq 1 \ : t - mh < \tau_i \leq t\big\} \ \in \ \{0, \ldots, m\}, \tag{3.2}$$

$$p(t,\alpha) \quad = \quad (\tau_{i+\iota(t,\alpha)}, \xi_{i+\iota(t,\alpha)})_{1\leq i\leq k(t,\alpha)} \ \in \ P_t(k(t,\alpha)). \tag{3.3}$$

Theorem 3.1. *The value functions satisfy the dynamic programming principle : for all k = 0, …, m, $(t,x,p) \in \mathcal{D}_k$,*

$$\begin{aligned} v_k(t,x,p) &= \sup_{\alpha\in\mathcal{A}_{t,p}} \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha})ds + \sum_{\tau_i+mh\le\theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-},\xi_i) \\ &\qquad + v_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha))\Big], \end{aligned} \tag{3.4}$$

where θ is any stopping time valued in $[t,T]$, possibly depending on α in (3.4). This means (i) for all $\alpha \in \mathcal{A}_{t,p}$, for all θ stopping time valued in $[t,T]$,

$$\begin{aligned} v_k(t,x,p) &\ge \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha})ds + \sum_{t<\tau_i+mh\le\theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-},\xi_i) \\ &\qquad + v_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha))\Big]. \end{aligned} \tag{3.5}$$

(ii) for all $\varepsilon > 0$, there exists $\alpha \in \mathcal{A}_{t,p}$ such that for all θ stopping time valued in $[t,T]$,

$$\begin{aligned} v_k(t,x,p) - \varepsilon &\le \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha})ds + \sum_{t<\tau_i+mh\le\theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-},\xi_i) \\ &\qquad + v_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha))\Big]. \end{aligned} \tag{3.6}$$

We now give an explicit consequence of the above dynamic programming that will be useful in the derivation of the corresponding analytic characterization. We introduce some additional notations. For all $t \in [0,T]$, we denote by $\mathcal{I}_t$ the set of pairs (τ,ξ) where τ is a stopping time, $t \le \tau \le T - mh$ or $\tau = \infty$ a.s., and ξ is a $\mathcal{F}_\tau$-measurable random variable valued in E. For any $p = (t_i,e_i)_{1\le i\le k} \in P_t(k)$, we denote $p_- = (t_i,e_i)_{2\le i\le k}$ with the convention that $p_- = \emptyset$ when $k = 1$.

When no impulse control is applied to the system, we denote by $X_s^{t,x,0}$ the solution to (2.1) with initial data $X_t = x$, and by $\mathcal{L}$ the associated infinitesimal generator :

$$\mathcal{L}\varphi = b(x).D_x\varphi + \frac{1}{2}\mathrm{tr}(\sigma\sigma'(x)D_x^2\varphi).$$

If $t \le T - mh$, we partition, for $k \in \{1,\dots,m\}$, the set $P_t(k)$ into $P_t(k) = P_t^1(k) \cup P_t^2(k)$ where

$$\begin{aligned} P_t^1(k) &= \Big\{p = (t_i,e_i)_{1\le i\le k} \in P_t(k) \ : t_k > t - h\Big\} \\ P_t^2(k) &= \Big\{p = (t_i,e_i)_{1\le i\le k} \in P_t(k) \ : t_k \le t - h\Big\}. \end{aligned}$$

Else if $t \ge T - mh$, we denote $P_t^1(k) = P_t(k)$ and $P_t^2(k) = \emptyset$. We easily see from the lag constraint on the pending orders that $P_t^2(k) = \emptyset$ if $k = m$, and so $P_t(m) = P_t^1(m)$.

Corollary 3.1. *Let $(t,x) \in [0,T) \times \mathbb{R}^d$.*
(1) For $k \in \{1,\dots,m\}$, and $p = (t_i,e_i)_{1\le i\le k} \in P_t^1(k)$, we have for any stopping time θ valued in $[t,(t_k+h)\wedge(t_1+mh))$:

$$v_k(t,x,p) = \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,0})ds + v_k(\theta, X_\theta^{t,x,0}, p)\Big]. \tag{3.7}$$

(2) For $k \in \{0, \ldots, m-1\}$, and $p = (t_i, e_i)_{1 \le i \le k} \in P_t^2(k)$, with the convention that $P_t^2(k) = \emptyset$ and $t_1 + mh = T$ when $k = 0$, we have for any stopping time θ valued in $[t, (t_1+mh) \wedge (t+h))$:

$$v_k(t,x,p) = \sup_{(\tau,\xi)\in\mathcal{I}_t} \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,0})ds + v_k(\theta, X_\theta^{t,x,0}, p)1_{\theta<\tau} + v_{k+1}(\theta, X_\theta^{t,x,0}, p \cup (\tau,\xi))1_{\tau\le\theta}\Big], \tag{3.8}$$

Interpretation and remarks

(1) $P_t^1(k)$ represents the set of k pending orders where the last order is within the period $(t-h, t]$ of nonintervention before t. Hence, from time t and until time $(t_k+h) \wedge (t_1+mh)$, we cannot intervene on the diffusion system and no pending order will be executed during this time period. This is mathematically formalized by relation (3.7).
(2) $P_t^2(k)$ represents the set of k pending orders where the last order is out of the period of nonintervention before t. Hence, at time t, one has two possible decisions : either one lets continue the system or one immediately intervene. In this latter case, this order adds to the previous ones. The mathematical formalization of these two choices is translated into relation (3.8).

In the next sections, we show how one can exploit these dynamic programming relations in order to characterize analytically the value functions by means of partial differential equations.

4 PDE system viscosity characterization

For $k = 1, \ldots, m$, let us introduce the subspace Θ_k of $[0, T-mh]^k$:

$$\Theta_k = \Big\{t^{(k)} = (t_i)_{1\le i\le k} \in [0, T-mh]^k \ : t_k - t_1 < mh, \ t_i - t_{i-1} \ge h, \ i = 2, \ldots, k\Big\}.$$

We shall write, by misuse of notation, $p = (t_i, e_i)_{1\le i\le k} = (t^{(k)}, e^{(k)})$, for any $t^{(k)} = (t_i)_{1\le i\le k} \in \Theta_k$, $e^{(k)} = (e_i)_{1\le i\le k} \in E^k$. By convention, we set $\Theta_k = E^k = \emptyset$ for $k = 0$. Notice that for all $t \in [0,T]$, and $p = (t^{(k)}, e^{(k)}) \in \Theta_k \times E^k$, $k = 0, \ldots, m$, we have

$$p \in P_t(k) \iff t \in \mathbb{T}_p(k),$$

where $\mathbb{T}_p(k)$ is the time domain in $[0,T]$ defined by :

$$\mathbb{T}_p(k) = [t_k, t_1 + mh).$$

By convention, we set $\mathbb{T}_p(k) = [0,T)$ for $k = 0$. We can then rewrite the domain $\mathcal{D}_k$ of the value function v_k in terms of union of time-space domains :

$$\mathcal{D}_k = \Big\{(t,x,p) \ : (t,x) \in \mathbb{T}_p(k) \times \mathbb{R}^d, \ p \in \Theta_k \times E^k\Big\}.$$

Therefore, the determination of the value function v_k, $k = 0, \ldots, m$, is equivalent to the determination of the function $v_k(.,.,p)$ on $\mathbb{T}_p(k) \times \mathbb{R}^d$ for all $p \in \Theta_k \times E_k$. The main goal of this paper is to provide an analytic characterization of these functions by means of the dynamic programming principle stated in the previous section.

For $k=0$, we set $\mathcal{D}_0 = [0,T)\times\mathbb{R}^d$. For $p=(t_i,e_i)_{1\le i\le k}\in\Theta_k\times E^k$, we partition the time domain $\mathbb{T}_p(k)$ into $\mathbb{T}_p(k)=\mathbb{T}_p^1(k)\cup\mathbb{T}_p^2(k)$ where

$$\begin{aligned}\mathbb{T}_p^2(k) &= \Big\{t\in\mathbb{T}_p(k)\cap[0,T-mh]\ : t\ge t_k+h\Big\} = [t_k+h,t_1+mh)\cap[0,T-mh],\\ \mathbb{T}_p^1(k) &= \mathbb{T}_p(k)\setminus\mathbb{T}_p^2(k)\end{aligned}$$

with the convention that $[s,t)=\emptyset$ if $s\ge t$. We then partition $\mathcal{D}_k$ into $\mathcal{D}_k=\mathcal{D}_k^1\cup\mathcal{D}_k^2$ where

$$\begin{aligned}\mathcal{D}_k^1 &= \Big\{(t,x,p)\in\mathcal{D}_k\ : t\in\mathbb{T}_p^1(k)\Big\}\\ \mathcal{D}_k^2 &= \Big\{(t,x,p)\in\mathcal{D}_k\ : t\in\mathbb{T}_p^2(k)\Big\}.\end{aligned}$$

Notice that for $k=1,\ldots,m$, and any $p\in\Theta_k\times E^k$, $\mathbb{T}_p^1(k)$ is never empty. In particular, $\mathcal{D}_k^1\neq\emptyset$. For $k=m$, and any $p=(t_i,e_i)_{1\le i\le m}\in\Theta_m\times E^m$, we have $t_m+h\ge t_1+mh$, and so $\mathbb{T}_p^2(m)=\emptyset$. Hence, $\mathcal{D}_m^2=\emptyset$ and $\mathcal{D}_m=\mathcal{D}_m^1$.

The PDE system to our control problem is formally derived by sending θ to $t<t_1+mh$ into dynamic programming relations (3.7)-(3.8). This provides equations for the value functions v_k on $\mathcal{D}_k$, which take the following nonstandard form, and are divided into :

$$-\frac{\partial v_k}{\partial t}(t,x,p)-\mathcal{L}v_k(t,x,p)-f(x) = 0 \quad \text{on} \quad \mathcal{D}_k^1, \quad k=0,\ldots,m, \tag{4.1}$$

$$\begin{aligned}\min\Big\{ &-\frac{\partial v_k}{\partial t}(t,x,p)-\mathcal{L}v_k(t,x,p)-f(x)\ ,\\ &v_k(t,x,p)-\sup_{e\in E}v_{k+1}(t,x,p\cup(t,e))\Big\} = 0 \quad \text{on} \quad \mathcal{D}_k^2,\ \ k=0,\ldots,m-1,\end{aligned} \tag{4.2}$$

with the convention that $\mathcal{D}_0^1=(T-mh,T)\times\mathbb{R}^d$ and $\mathcal{D}_0^2=[0,T-mh]\times\mathbb{R}^d$.

As usual, the value functions need not be smooth, and even not known to be continuous a priori, and we shall work with the notion of (discontinuous) viscosity solutions (see [27] or [38] for classical references on the subject), which we adapt in our context as follows. For a locally bounded function w_k on $\mathcal{D}_k$, we denote $\underline{w_k}$ (resp. $\overline{w}_k$) its lower semicontinuous (resp. upper-semicontinuous) envelope, i.e.

$$\begin{aligned}\underline{w_k}(t,x,p) &= \liminf_{(t',x',p')\to(t,x,p)} w_k(t',x',p'),\\ \overline{w_k}(t,x,p) &= \limsup_{(t',x',p')\to(t,x,p)} w_k(t',x',p'), \quad (t,x,p)\in\mathcal{D}_k,\ k=0,\ldots,m.\end{aligned}$$

Definition 4.1. *We say that a family of locally bounded functions w_k on $\mathcal{D}_k$, $k=0,\ldots,m$, is a viscosity supersolution (resp. subsolution) of* (4.1)-(4.2) *on $\mathcal{D}_k$, $k=0,\ldots,m$, if :*
(i) for all $k=1,\ldots,m$, $(t_0,x_0,p_0)\in\mathcal{D}_k^1$, and $\varphi\in C^2(\mathcal{D}_k^1)$, which realizes a local minimum of $\underline{w_k}-\varphi$ (resp. maximum of $\overline{w_k}-\varphi$), we have

$$-\frac{\partial\varphi}{\partial t}(t_0,x_0,p_0)-\mathcal{L}\varphi(t_0,x_0)-f(x_0) \ \ge\ 0 \quad (\textit{resp.}\ \le 0).$$

(ii) for all $k = 0, \ldots, m-1$, $(t_0, x_0, p_0) \in \mathcal{D}_k^2$, and $\varphi \in C^2(\mathcal{D}_k^2)$, which realizes a local minimum of $\underline{w_k} - \varphi$ (resp. maximum of $\overline{w_k} - \varphi$), we have

$$\begin{aligned}\min\Big\{ -\frac{\partial\varphi}{\partial t}(t_0, x_0, p_0) - \mathcal{L}\varphi(t_0, x_0, p_0) - f(x_0)\ , \\ \underline{w_k}(t_0, x_0, p_0) - \sup_{e\in E}\underline{w_{k+1}}(t_0, x_0, p_0 \cup (t_0, e))\Big\} \quad \geq \quad 0\end{aligned}$$

(resp.

$$\begin{aligned}\min\Big\{ -\frac{\partial\varphi}{\partial t}(t_0, x_0, p_0) - \mathcal{L}\varphi(t_0, x_0, p_0) - f(x_0)\ , \\ \overline{w_k}(t_0, x_0, p_0) - \sup_{e\in E}\overline{w_{k+1}}(t_0, x_0, p_0 \cup (t_0, e))\Big\} \quad \leq \quad 0)\end{aligned}$$

We say that a family of locally bounded functions w_k on $\mathcal{D}_k$, $k = 0, \ldots, m$, is a viscosity solution of (4.1)-(4.2) *if it is a viscosity supersolution and subsolution of* (4.1)-(4.2)*.*

We then state the viscosity property of the value functions to our control problem.

Proposition 4.1. *(Viscosity property)*
The family of value functions v_k, $k = 0, \ldots, m$, is a viscosity solution to (4.1)-(4.2)*. Moreover, for all $k = 0, \ldots, m-1$, $(t, x, p) \in \mathcal{D}_k^2$, $p = (t_i, e_i)_{1\leq i\leq k}$ with $t = t_k + h$ or $t = T - mh$, we have :*

$$\underline{v_k}(t, x, p) \geq \sup_{e\in E}\underline{v_{k+1}}(t, x, p \cup (t, e)), \tag{4.3}$$

In order to have a complete characterization of the value functions, and so of our control problem, we need to determine the suitable boundary conditions. These concern for $k = 1, \ldots, m$ the time-boundary of $\mathcal{D}_k$, i.e. the points $(t_1 + mh, x, p)$ for $x \in \mathbb{R}^d$, $p = (t_i, e_i)_{1\leq i\leq k}$ $\in \Theta_k \times E^k$, and also the value function v_0 on (T, x), $x \in \mathbb{R}^d$. For a locally bounded function w_k on $\mathcal{D}_k$, $k = 1, \ldots, m$, we denote

$$\begin{aligned}\overline{w_k}(t_1 + mh, x, p) &= \limsup_{\substack{(t,x',p')\to(t_1+mh,x,p)\\(t,x',p')\in\mathcal{D}_k}} w_k(t, x', p'), \\ \underline{w_k}(t_1 + mh, x, p) &= \liminf_{\substack{(t,x',p')\to(t_1+mh,x,p)\\(t,x',p')\in\mathcal{D}_k}} v_k(t, x', p'),\ \ x \in \mathbb{R}^d,\ p = (t_i, e_i)_{1\leq i\leq k} \in \Theta_k,\end{aligned}$$

and if these two limits are equal, we set

$$w_k((t_1 + mh)^-, x, p) \ = \ \overline{w_k}(t_1 + mh, x, p) \ = \ \underline{w_k}(t_1 + mh, x, p).$$

Proposition 4.2. *(Boundary data)*
(i) For $k = 1, \ldots, m$, $p = (t_i, e_i)_{1\leq i\leq k} \in \Theta_k \times E^k$, $x \in \mathbb{R}^d$, $v_k((t_1 + mh)^-, x, p)$ exists and :

$$v_k((t_1 + mh)^-, x, p) \ = \ c(x, e_1) + v_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-). \tag{4.4}$$

(ii) At time T, for all $x \in \mathbb{R}^d$, $v_0(T^-, x)$ exists and:

$$v_0(T^-, x) = g(x) \tag{4.5}$$

We can now state the unique PDE characterization result for our control delay problem.

Theorem 4.1. *The family of value functions v_k, $k = 0, \ldots, m$, is the unique viscosity solution to* (4.1)-(4.2)*, which satisfy* (4.3)*, the boundary data* (4.4)-(4.5)*, and the linear growth condition* (2.9)*. Moreover, v_k is continuous on $\mathcal{D}_k$, $k = 0, \ldots, m$.*

Remark 4.1. (Case $m = 1$)
In the particular case where the execution delay is equal to the intervention lag, i.e. $m = 1$, we have two value functions v_0 and v_1, and the system (4.1)-(4.2) may be significantly simplified. First, remark that linear equation (4.1) on $\mathbb{R}^d \times (T - mh, T]$ together with terminal condition (4.5) leads to:

$$v_0(t,x) \;=\; \mathbb{E}\Big[\int_t^T f(X_s^{t,x,0})ds + g(X_T^{t,x,0})\Big], \;\; (t,x) \in (T-mh, T]. \tag{4.6}$$

Furthermore, from the linear PDE (4.1) and the boundary data (4.4) for $k = m = 1$, we have the Feynman-Kac representation :

$$v_1(t,x,(t_1,e_1)) \;=\; \mathbb{E}\Big[\int_t^{t_1+h} f(X_s^{t,x,0})ds + c(X_{t_1+h}^{t,x,0}, e) + v_0(t_1+h, \Gamma(X_{t_1+h}^{t,x,0}, e))\Big], \tag{4.7}$$

for all $(t_1, e_1) \in [0, T-h] \times E$, $(t,x) \in [t_1, t_1+h) \times \mathbb{R}^d$. By plugging (4.7) for $t = t_1$ into (4.2) for $k = 0$, we obtain the variational inequality satisfied by v_0 :

$$\begin{aligned} 0 \;=\; \min\Big\{ &-\frac{\partial v_0}{\partial t} - \mathcal{L}v_0 - f\,, \\ v_0 - \sup_{e\in E}\mathbb{E}\Big[\int_t^{t+h} &f(X_s^{t,x,0})ds + c(X_{t+h}^{t,x,0}, e) + v_0(t+h, \Gamma(X_{t+h}^{t,x,0}, e))\Big]\Big\} \quad \text{on} \quad [0, T-h] \times \mathbb{R}^d, \end{aligned} \tag{4.8}$$

together with the terminal condition for $k = 0$ (see (4.6)). Therefore, in the case $m = 1$, and as observed in [60], the original problem is reduced to a no-delay impulse control problem (4.8) for v_0, and v_1 is explicitly related to v_0 by (4.7). Equations (4.8)-(4.6) can be solved by iterated optimal stopping problems, see the details in the next chapter in the more general case $m \geq 1$.

Remark 4.2. In the general case $m \geq 1$, we point out the peculiarities of the PDE characterization for our control delay problem.
1. The dynamic programming coupled system (4.1)-(4.2) has a nonstandard form. For fixed k, there is a discontinuity on the differential operator of the equation satisfied by v_k on $\mathcal{D}_k$. Indeed, the PDE is divided into a linear equation on the subdomain $\mathcal{D}_k^1$, and a variational inequality with obstacle involving the value function v_{k+1} on the subdomain $\mathcal{D}_k^2$. Moreover, the time domain $\mathbb{T}_p(k)$ of $\mathcal{D}_k$ for $v_k(., x, p)$ depends on the argument $p \in \Theta_k$. With respect to usual comparison principle of nonlinear PDE, we state an uniqueness result for viscosity solutions satisfying in addition the inequality (4.3) at the discontinuity of the differential operator.
2. The boundary data also present some specificities. For fixed k, the condition in (4.4) concerns as usual data on the time-boundary of the domain $\mathcal{D}_k$ on which the value function v_k satisfies a PDE. However, it involves data on the value function v_{k-1}, which is a priori not known.

3. The continuity property of the value functions v_k on $\mathcal{D}_k$ is not at all obvious a priori from the very definitions of v_k, and is proved actually as consequences of comparison principles and boundary data for the system (4.1)-(4.2), see Proposition 6.4. In particular, if assumption (2.7) is relaxed, then continuity does not hold necessarily for the value function. For example, by taking $b = \sigma = f = g = 0$ and $c(x, e) = 1$ for all $(x, e) \in \mathbb{R}^d \times E$, we easily see that $v_0(t, x) = \max(0, \lceil \frac{T-mh-t}{h} \rceil)$, where $\lceil x \rceil$ denotes the smallest integer which is superior to x, which is obviously not continuous.

The PDE characterization in Theorem 4.1 means that the value functions are in theory completely determined by the resolution of the PDE system (4.1)-(4.2) together with the boundary data (4.4)-(4.5). We show in the next section how to solve this system and compute in practice these value functions and the associated optimal impulse controls.

5 Description of the optimal impulse control

In view of the above dynamic programming relations, and the general theory of optimal stopping (see [34]), we can describe the structure of the optimal impulse control for $V_0 = v_0(0, X_0)$ in terms of the value functions. Let us define the following quantities :

▶ **Initialization :** $n = 0$

- *given an initial pending order number* $k = 0$, we define

$$
\begin{aligned}
\tilde{\tau}_1^{(0)} &= \inf\left\{t \geq 0 \ : v_0(t, X_t^{\alpha^*}) \ = \ \sup_{e \in E} v_1(t, X_t^{\alpha^*}, (t, e))\right\} \wedge T, \\
\tilde{e}_1^{(0)} &\in \arg\max_{e \in E} v_1(\tilde{\tau}_1^{(0)}, X^{\alpha^*}_{\tilde{\tau}_1^{(0)}}, (\tilde{\tau}_1^{(0)}, e)).
\end{aligned}
$$

  If $\tilde{\tau}_1^{(0)} + mh > T$, we stop the induction at $n = 0$, otherwise continue to the next item :

- *Pending orders number* $k \to k+1$ (this step is empty when $m = 1$) from $k = 1$:

$$
\begin{aligned}
\tilde{\tau}_{k+1}^{(0)} &= \inf\Big\{t \geq \tilde{\tau}_k^{(0)} + h \ : \\
&\qquad v_k(t, X_t^{\alpha^*}) \ = \ \sup_{e \in E} v_{k+1}(t, X_t^{\alpha^*}, (\tilde{\tau}_i^{(0)}, \tilde{e}_i^{(0)})_{1 \leq i \leq k} \cup (t, e))\Big\} \wedge T, \\
\tilde{e}_{k+1}^{(0)} &\in \arg\max_{e \in E} v_{k+1}(\tilde{\tau}_{k+1}^{(0)}, X^{\alpha^*}_{\tilde{\tau}_{k+1}^{(0)}}, (\tilde{\tau}_i^{(n)}, \tilde{e}_i^{(0)})_{1 \leq i \leq k} \cup (\tilde{\tau}_{k+1}^{(0)}, e)).
\end{aligned}
$$

  As long as $\tilde{\tau}_k^{(0)} \leq \tilde{\tau}_1^{(0)} + mh$, increment $k \to k+1 : \tilde{\tau}_k^{(0)} \to \tilde{\tau}_{k+1}^{(0)}$, until

$$
k_0 \ = \ \sup\left\{k \ : \tilde{\tau}_k^{(0)} \leq \tilde{\tau}_1^{(0)} + mh\right\} \ \in \ \{1, \ldots, m\},
$$

  and increment the induction on n by the following step :

▶ $n \to n+1$ **:**

- *given an initial pending orders number* $k = k_n - 1$, we define

$$\begin{aligned} \tilde{\tau}_{k_n}^{(n+1)} &= \inf\big\{t \geq (\tilde{\tau}_1^{(n)} + mh) \vee (\tilde{\tau}_{k_n}^{(n)} + h) \ : \\ &\qquad v_{k_n-1}(t, X_t^{\alpha^*}, \tilde{p}_{n^-}) \ = \ \sup_{e \in E} v_{k_n}(t, X_t^{\alpha^*}, \tilde{p}_{n^-} \cup (t,e))\big\} \wedge T, \\ \tilde{e}_{k_n}^{(n+1)} &\in \arg\max_{e \in E} v_{k_n}(\tilde{\tau}_1^{n+1}, X_{\tilde{\tau}_1^{n+1}}^{\alpha^*}, \tilde{p}_{n^-} \cup (\tilde{\tau}_{k_n}^{n+1}, e)), \end{aligned}$$

where we set $\tilde{p}_{n^-} = (\tilde{\tau}_i^{(n)}, \tilde{e}_i^{(n)})_{2 \leq i \leq k_n}$. We denote $\tilde{\tau}_1^{(n+1)} = \tilde{\tau}_2^{(n)}$ if $k_n > 1$, and $\tilde{\tau}_1^{(n+1)}$ $= \tilde{\tau}_{k_n}^{n+1}$ if $k_n = 1$. If $\tilde{\tau}_1^{(n+1)} + mh > T$, we stop the induction at $n+1$, otherwise continue to the next item :

- *Pending orders number* $k \to k+1$ (this step is empty when $m = 1$) from $k = k_n$:

$$\begin{aligned} \tilde{\tau}_{k+1}^{(n+1)} &= \inf\big\{t \geq \tilde{\tau}_k^{(n+1)} + h \ : \\ &\qquad v_k(t, X_t^{\alpha^*}) \ = \ \sup_{e \in E} v_{k+1}(t, X_t^{\alpha^*}, \tilde{p}_{n^-} \cup (\tilde{\tau}_i^{(n+1)}, \tilde{e}_i^{(n+1)})_{k_n \leq i \leq k} \cup (t,e))\big\} \wedge T \\ \tilde{e}_{k+1}^{(n+1)} &\in \arg\max_{e \in E} v_{k+1}(\tilde{\tau}_{k+1}^{(n+1)}, X_{\tilde{\tau}_{k+1}^{(n+1)}}^{\alpha^*}, \tilde{p}_{n^-} \cup (\tilde{\tau}_i^{(n+1)}, \tilde{e}_i^{(n+1)})_{k_n \leq i \leq k} \cup (\tilde{\tau}_{k+1}^{(n+1)}, e)) \end{aligned}$$

As long as $\tilde{\tau}_k^{(n+1)} \leq \tilde{\tau}_1^{(n+1)} + mh$, increment $k \to k+1$: $\tilde{\tau}_k^{(n+1)} \to \tilde{\tau}_{k+1}^{(n+1)}$, until

$$k_{n+1} \ = \ \sup\big\{k \ : \tilde{\tau}_k^{(n+1)} \leq \tilde{\tau}_1^{(n+1)} + mh\big\} \ \in \ \{1, \ldots, m\},$$

and continue the induction on n : $n \to n+1$ until $\tilde{\tau}_1^{(n+1)} + mh > T$.

The optimal impulse control is given by the finite sequence $\{(\tilde{\tau}_k^{(n)}, \tilde{e}_k^{(n)})_{k_{n-1} \leq k \leq k_n}, n = 0, \ldots, N\}$, where $N = \inf\{n \geq 0 \ : \tilde{\tau}_1^{(n)} + mh > T\}$, and we set by convention $k_{-1} = 1$.

6 Proofs of main results

6.1 Dynamic programming principle

From the dynamics (2.3) of the controlled process, we derive easily the following properties (recall the notations (3.1)-(3.2)-(3.3)) :
• Markov property of the pair $(X^\alpha, p(., \alpha))$ for any $\alpha \in \mathcal{A}$, in the sense that

$$\mathbb{E}\Big[\varphi(X_{\theta_2}^\alpha)\Big|\mathcal{F}_{\theta_1}\Big] \ = \ \mathbb{E}\Big[\varphi(X_{\theta_2}^\alpha)\Big|(X_{\theta_1}^\alpha, p(\theta_1, \alpha))\Big],$$

for any bounded measurable function φ, and stopping times $\theta_1 \leq \theta_2$ a.s.
• Causality of the control, in the sense that for any $\alpha = (\tau_i, \xi_i)_{i \geq 1} \in \mathcal{A}$, and θ stopping time,

$$\alpha^\theta \ \in \ \mathcal{A}_{\theta, p(\theta, \alpha)}, \quad \text{and} \quad p(\theta, \alpha) \ \in \ k(\theta, \alpha) \quad a.s.$$

where we set $\alpha^\theta = (\tau_{i+\iota(\theta,\alpha)}, \xi_{i+\iota(\theta,\alpha)})_{i \geq 1}$.

• Pathwise uniqueness of the state process,

$$X^{t,x,p,\alpha} = X^{\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha), \alpha^\theta} \quad \text{on} \quad [\theta, T],$$

for any $(t,x,p) \in \mathcal{D}_k$, $k = 0, \ldots, m$, $\alpha \in \mathcal{A}_{t,p}$, and $\theta \in \mathcal{T}_{t,T}$ the set of stopping times valued in $[t,T]$.

From the above properties, we deduce by usual arguments the inequality (3.6) of the dynamic programming principle, which can be formulated equivalently in

Proposition 6.1. *For all $k = 0, \ldots, m$, $(t,x,p) \in \mathcal{D}_k$, we have*

$$\begin{aligned} v_k(t,x,p) \quad \leq \quad & \sup_{\alpha \in \mathcal{A}_{t,p}} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha}) ds + \sum_{t < \tau_i + mh \leq \theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-}, \xi_i) \\ & \qquad + v_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha)) \Big]. \end{aligned}$$

Proof. Fix $(t,x,p) \in \mathcal{D}_k$, $k = 0, \ldots, m$, and take arbitrary $\alpha \in \mathcal{A}_{t,p}$, $\theta \in \mathcal{T}_{t,T}$. From the definitions of the performance criterion and the value functions, the law of iterated conditional expectations, Markov property, pathwise uniqueness, and causality features of our model, we get the successive relations

$$\begin{aligned} J_k(t,x,p,\alpha) \quad = \quad & \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha}) ds + \sum_{t < \tau_i + mh \leq \theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-}, \xi_i) \\ & + \mathbb{E}\Big[\int_\theta^T f(X_s^{t,x,p,\alpha}) ds + g(X_T^{t,x,p,\alpha}) + \sum_{\theta < \tau_i + mh \leq T} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-}, \xi_i) \Big| \mathcal{F}_\theta \Big]\Big] \\ = \quad & \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha}) ds + \sum_{t < \tau_i + mh \leq \theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-}, \xi_i) \\ & + J_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha), \alpha^\theta) \Big] \\ \leq \quad & \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha}) ds + \sum_{t < \tau_i + mh \leq \theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-}, \xi_i) \\ & + v_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha)) \Big]. \end{aligned}$$

Since θ and α are arbitrary, we obtain the required inequality. □

As usual, the inequality (3.5) of the dynamic programming principle requires in addition to the Markov, causality and pathwise uniqueness properties, a measurable selection theorem. This inequality can be formulated equivalently in

Proposition 6.2. *For all $k = 0, \ldots, m$, $(t,x,p) \in \mathcal{D}_k$, we have*

$$\begin{aligned} v_k(t,x,p) \quad \geq \quad & \sup_{\alpha \in \mathcal{A}_{t,p}} \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha}) ds + \sum_{t < \tau_i + mh \leq \theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-}, \xi_i) \\ & \qquad + v_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha)) \Big]. \end{aligned}$$

Proof. Fix $(t,x,p) \in \mathcal{D}_k$, $k = 0,\ldots,m$, and arbitrary $\alpha \in \mathcal{A}_{t,p}$, $\theta \in \mathcal{T}_{t,T}$. By definition of the value functions, for any $\varepsilon > 0$ and $\omega \in \Omega$, there exists $\alpha_{\varepsilon,\omega} \in \mathcal{A}_{\theta(\omega),p(\theta(\omega),\alpha(\omega))}$, which is an ε-optimal control for $v_{k(\theta(\omega),\alpha(\omega))}$ at $(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha))(\omega)$. By a measurable selection theorem (see e.g. Chapter 7 in [15]), there exists $\bar\alpha_\varepsilon \in \mathcal{A}_{\theta,p(\theta,\alpha)}$ s.t. $\bar\alpha_\varepsilon(\omega) = \alpha_{\varepsilon,\omega}(\omega)$ a.s., and so

$$v_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha)) - \varepsilon \quad \leq \quad J_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha), \bar\alpha_\varepsilon) \qquad a.s. \tag{6.1}$$

Now, we define by concatenation the impulse control $\bar\alpha$ consisting of the impulse control components of α until (including eventually) time τ, and the impulse control components of $\bar\alpha_\varepsilon$ strictly after time τ. By construction, $\bar\alpha \in \mathcal{A}_{t,p}$, $X^{t,x,p,\bar\alpha} = X^{t,x,p,\alpha}$ on $[t,\theta]$, $k(\theta,\bar\alpha) = k(\theta,\alpha)$, $p(\theta,\bar\alpha) = p(\theta,\alpha)$, and $\bar\alpha^\theta = \bar\alpha_\varepsilon$. Hence, similarly as in Proposition 6.1, by using law of iterated conditional expectations, Markov property, pathwise uniqueness, and causality features of our model, we get

$$\begin{aligned} J_k(t,x,p,\bar\alpha) \quad = \quad & \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha})ds + \sum_{t<\tau_i+mh\leq\theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-}, \xi_i) \\ & \quad + J_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha), \bar\alpha_\varepsilon)\Big]. \end{aligned}$$

Together with (6.1), this implies

$$\begin{aligned} v_k(t,x,p) \ \geq \ J_k(t,x,p,\bar\alpha) \quad \geq \quad & \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,p,\alpha})ds + \sum_{t<\tau_i+mh\leq\theta} c(X^{t,x,p,\alpha}_{(\tau_i+mh)^-}, \xi_i) \\ & \quad + v_{k(\theta,\alpha)}(\theta, X_\theta^{t,x,p,\alpha}, p(\theta,\alpha))\Big] - \varepsilon. \end{aligned}$$

From the arbitrariness of ε, α, and θ, this proves the required result. □

We end this paragraph by proving Corollary 3.1.

Proof of Corollary 3.1.
(i) Fix $k \in \{1,\ldots,m\}$, $(t,x) \in [0,T] \times \mathbb{R}^d$, $p = (t_i,e_i)_{1\leq i\leq k} \in P_t^1(k)$ such that $t_1 + mh \leq T$, and θ stopping time valued in $[t,(t_k+h)\wedge(t_1+mh))$. Then, we observe that for all $\alpha = (\tau_i,\xi_i)_{i\geq1} \in \mathcal{A}_{t,p}$, $X^{t,x,p,\alpha} = X^{t,x,0}$ on $[t,\theta]$, $\tau_i + mh > \theta$, $k(\theta,\alpha) = k$, and $p(\theta,\alpha) = p$ a.s. Hence, relation (3.7) follows immediately from (3.4).

(ii) For $k \in \{0,\ldots,m-1\}$, $p = (t_i,e_i)_{1\leq i\leq k} \in P_t^2(k)$ such that $t_1+mh \leq T$, and θ stopping time valued in $[t,(t_1+mh)\wedge(t+h))$. Let $\alpha = (\tau_i,\xi_i)_{i\geq1}$ be some arbitrary element in $\mathcal{A}_{t,p}$, and set $\tau = \tau_{k+1}$, $\xi = \xi_{k+1}$. Notice that $(\tau,\xi) \in \mathcal{I}_t$. Then, we see that $X^{t,x,p,\alpha} = X^{t,x,0}$ on $[t,\theta]$, $\tau_i + mh > \theta$, $k(\theta,\alpha) = k$, $p(\theta,\alpha) = p$ if $\theta < \tau$, and $k(\theta,\alpha) = k+1$, $p(\theta,\alpha) = p \cup (\tau,\xi)$ if $\theta \geq \tau$. We deduce from (3.5) that

$$\begin{aligned} v_k(t,x,p) \quad \geq \quad & \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,0})ds + v_k(\theta, X_\theta^{t,x,0}, p)1_{\theta<\tau} \\ & \quad + v_{k+1}(\theta, X_\theta^{t,x,0}, p \cup (\tau,\xi))1_{\tau\leq\theta}\Big], \end{aligned}$$

and this inequality holds for any $(\tau,\xi) \in \mathcal{I}_t$ by arbitrariness of α. Furthermore, from (3.6), for all $\varepsilon > 0$, there exists $(\tau, \xi) \in \mathcal{I}_t$ s.t.

$$\begin{aligned} v_k(t,x,p) - \varepsilon \quad \leq \quad & \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,0}) ds + v_k(\theta, X_\theta^{t,x,0}, p) 1_{\theta<\tau} \\ & \qquad + v_{k+1}(\theta, X_\theta^{t,x,0}, p \cup (\tau,\xi)) 1_{\tau \leq \theta} \Big]. \end{aligned}$$

The two previous inequalities give the required relation

$$\begin{aligned} v_k(t,x,p) \quad = \quad & \sup_{(\tau,\xi)\in\mathcal{I}_t} \mathbb{E}\Big[\int_t^\theta f(X_s^{t,x,0}) ds + v_k(\theta, X_\theta^{t,x,0}, p) 1_{\theta<\tau} \\ & \qquad + v_{k+1}(\theta, X_\theta^{t,x,0}, p \cup (\tau,\xi)) 1_{\tau \leq \theta} \Big]. \end{aligned}$$

6.2 Viscosity properties

In this paragraph, we prove the viscosity property stated in Proposition 4.1. We first state an auxiliary result. For any locally bounded function u on $\mathcal{D}_{k+1}$, $k = 0, \ldots, m-1$, we define the locally bounded function $\mathcal{H}u$ on $\mathcal{D}_k^2$ by $\mathcal{H}u(t,x,p) = \sup_{e\in E} u(t,x,p\cup(t,e))$.

Lemma 6.1. *Let u be a locally bounded function on $\mathcal{D}_{k+1}$, $k = 0, \ldots, m-1$. Then, $\mathcal{H}\overline{u}$ is upper-semicontinuous, and $\overline{\mathcal{H}u} \leq \mathcal{H}\overline{u}$.*

Proof. Fix some $(t,x,p) \in \mathcal{D}_k^2$, and let $(t_n,x_n,p_n)_{n\geq 1}$ be a sequence in $\mathcal{D}_k^2$ converging to (t,x,p) as n goes to infinity. Since $\overline{u}$ is upper-semicontinuous, and E is compact, there exists a sequence $(e_n)_n$ valued in E, such that

$$\mathcal{H}\overline{u}(t_n,x_n,p_n) \quad = \quad \overline{u}(t_n,x_n,p_n \cup (t_n,e_n)), \quad n \geq 1.$$

The sequence $(e_n)_n$ converges, up to a subsequence, to some $\hat{e} \in E$, and so

$$\mathcal{H}\overline{u}(t,x,p) \ \geq \ \overline{u}(t,x,p\cup(t,\hat{e})) \ \geq \ \limsup_{n\to\infty} \overline{u}(t_n,x_n,p_n\cup(t_n,e_n)) \ = \ \limsup_{n\to\infty} \mathcal{H}\overline{u}(t_n,x_n,p_n),$$

which shows that $\mathcal{H}\overline{u}$ is upper-semicontinuous.

On the other hand, fix some $(t,x,p) \in \mathcal{D}_k^{2,m}$, and let $(t_n,x_n,p_n)_{n\geq 1}$ be a sequence in $\mathcal{D}_k^2$ converging to (t,x,p) s.t. $\mathcal{H}u(t_n,x_n,p_n)$ converges to $\overline{\mathcal{H}u}(t,x,p)$. Then, we have

$$\overline{\mathcal{H}u}(t,x,p) \ = \ \lim_{n\to\infty} \mathcal{H}u(t_n,x_n,p_n) \ \leq \ \limsup_{n\to\infty} \mathcal{H}\overline{u}(t_n,x_n,p_n) \ \leq \ \mathcal{H}\overline{u}(t,x,p),$$

which shows that $\overline{\mathcal{H}u} \leq \mathcal{H}\overline{u}$. □

Now, we prove the sub and supersolution property of the family v_k, $k = 0, \ldots, m$. There is no difficulty on the domain $\mathcal{D}_k^1$ since locally no impulse control is possible. Hence, in this case, the viscosity properties can be derived as for an uncontrolled state process, and the proof is standard from the dynamic programming principle (3.7), see e.g. [58]. Notice that since the domain $\mathbb{T}_p^1(k)$ is open in $\mathbb{T}_p(k)$, we have no problem at the boundary. Indeed, this set is open at $(t_k+h)\wedge(t_1+mh)$ and eventually $T-mh$, which is the usual situation, and the closedness at t_k and T does not introduce difficulties, as the value function is not defined before t_k and after T. Hence, when taking approximations of the upper and lower semicontinous envelopes of v_k, we only need to consider points of the domain such that $t \geq t_k$, where the dynamic programming relation (3.7) holds. The proof of the viscosity property of the value functions v_k to (4.2) on $\mathcal{D}_k^2$ is more subtle. Indeed, in addition to the specific form of equation (4.2), we have to carefully address the discontinuity of the PDE system (4.1)-(4.2) on the boudaries $t_k + h$ and eventually $T - mh$ of $\mathbb{T}_p^2(k)$. In the sequel, we focus on the domain $\mathcal{D}_k^2$, $k = 0, \ldots, m-1$.

Proof of the supersolution property on $\mathcal{D}_k^2$.

We first prove that for $k = 0, \ldots, m-1$, $(t_0, x_0, p_0) \in \mathcal{D}_k^2$:

$$\underline{v_k}(t_0, x_0, p_0) \geq \sup_{e \in E} \underline{v_{k+1}}(t_0, x_0, p_0 \cup (t_0, e)). \tag{6.2}$$

By definition of $\underline{v_k}$, there exists a sequence $(t_n, x_n, p_n)_{n \geq 1} \in \mathcal{D}_k^m$ such that :

$$v_k(t_n, x_n, p_n) \to \underline{v_k}(t_0, x_0, p_0) \quad \text{with} \quad (t_n, x_n, p_n) \to (t_0, x_0, p_0). \tag{6.3}$$

We set $p_0 = (t_i^0, e_i^0)_{1 \leq i \leq k}$, $p_n = (t_i^n, e_i^n)_{1 \leq i \leq k}$, and we distinguish the three following cases :

- If $t_k^0 + h < t_0 < T - mh$, then, for n sufficiently large, we have $t_k^n + h \leq t_n \leq T - mh$, i.e. $p_n \in P_{t_n}^2(k)$. Hence, from the dynamic programming principle by making an immediate impulse control, i.e. by applying (3.8) to $v_k(t_n, x_n, p_n)$ with $\theta = \tau = t_n$, and $e \in E$, we have

$$v_k(t_n, x_n, p_n) \quad \geq \quad v_{k+1}(t_n, x_n, p_n \cup (t_n, e)) \;\geq\; \underline{v_{k+1}}(t_n, x_n, p_n \cup (t_n, e)).$$

  By sending n to infinity with (6.3), and since $\underline{v_{k+1}}$ is lower-semicontinuous, we obtain the required relation (6.2) from the arbitrariness of e in E.

- if $t_0 = t_k^0 + h \neq T - mh$, we apply the dynamic programming principle by making an impulse control as soon as possible. This means that in relation (3.5) for $v_k(t_n, x_n, p_n)$, we choose $\alpha = (\tau_i, \xi_i)_{i \geq 1} \in \mathcal{A}_{t_n, p_n}$, $\theta = \tau_{k+1} = \theta_n := t_n \vee (t_k^n + h)$, $\xi_{k+1} = e \in E$, so that :

$$\begin{aligned} v_k(t_n, x_n, p_n) \quad &\geq \quad \mathbb{E}\Big[\int_{t_n}^{\theta_n} f(X_s^n)ds + \sum_{t_n < \tau_i + mh \leq \theta_n} c(X^n_{(\tau_i + mh)^-}, \xi_i) \\ &\qquad + \underline{v_{k+1}}(\theta_n, X^n_{\theta_n}, p_n \cup (\theta_n, e))\Big]. \end{aligned}$$

  Here $X^n := X^{t_n, x_n, 0}$. Since t_n, $\theta_n \to t_0$, $p_n \to p_0$, $X^n_{\theta_n} \to x_0$ a.s., as n goes to infinity, and from estimate (2.8) and the linear growth condition on f, c, $\underline{v_{k+1}}$, we can use the dominated convergence theorem to obtain :

$$\underline{v_k}(t_0, x_0, p_0) \geq \underline{v_{k+1}}(t_0, x_0, p_0 \cup (t_0, e)),$$

  which implies (6.2) from the arbitrariness of $e \in E$.

- if $t_0 = T - mh$, we show from condition (2.7) that it is not optimal to decide an impulse intervention. First, notice from the definition of the value function and from the constraints on the impulse controls that, for all $e \in E$:

$$\begin{aligned} v_{k+1}(t_0, x_0, p_0 \cup (t_0, e)) &= \mathbb{E}\Big[\int_{t_0}^T f(X_s^{t_0,x_0,p_0})ds + g\big(\Gamma(X_{T^-}^{t_0,x_0,p_0}, e)\big) \\ &+ \sum_{i=0}^k c(X_{(t_i^0+mh)^-}^{t_0,x_0,p_0}, e_i^0) + c(X_{T^-}^{t_0,x_0,p_0}, e)\Big]. \end{aligned} \quad (6.4)$$

Moreover, by definition of v_k, and by choosing not to decide an impulse intervention, we get for all n :

$$v_k(t_n, x_n, p_n) \;\geq\; \mathbb{E}\Big[\int_{t_n}^T f(X_s^{t_n,x_n,p_n})ds + g\big(X_{T^-}^{t_n,x_n,p_n}\big) + \sum_{i=0}^k c(X_{(t_i^n+mh)^-}^{t_n,x_n,p_n}, e_i^n)\Big].$$

Hence, by the continuity and the linear growth conditions of f, g, Γ, c together with the dominated convergence theorem, we get by sending n to infinity into the previous inequality :

$$\underline{v_k}(t_0, x_0, p_0) \;\geq\; \mathbb{E}\Big[\int_{t_0}^T f(X_s^{t_0,x_0,p_0})ds + g\big(X_{T^-}^{t_0,x_0,p_0}\big) + \sum_{i=0,..,k} c(X_{(t_{i,0}+mh)^-}^{t_0,x_0,p_0}, e_{i,0})\Big].$$

Finally, by using Assumption (2.7) and equality (6.4), we get :

$$\underline{v_k}(t_0, x_0, p_0) \;\geq\; v_{k+1}(t_0, x_0, p_0 \cup (t_0, e)) \;\geq\; \underline{v_{k+1}}(t_0, x_0, p_0 \cup (t_0, e)),$$

which proves the required inequality from the arbitrariness of e in E.

Finaly, in order to complete the viscosity supersolution property of v_k to (4.2) on $\mathcal{D}_k^2$, it remains to show that v_k is a supersolution to :

$$-\frac{\partial v_k}{\partial t}(t, x, p) - \mathcal{L}v_k(t, x, p) - f(x) \;\geq\; 0,$$

on $\mathcal{D}_k^2$. This proof is standard by using the dynamic programming relation (3.8) with $\tau = \infty$ and Itô's formula, see [58] for the details. □

Proof of the subsolution property on $\mathcal{D}_k^2$.

We follow arguments in [52]. Let $(t_0, x_0, p_0) \in \mathcal{D}_k^{2,m}$ and $\varphi \in C^{1,2}(\mathcal{D}_k^2)$ such that $\overline{v_k}(t_0, x_0, p_0) = \varphi(t_0, x_0, p_0)$ and $\varphi \geq \overline{v_k}$ on $\mathcal{D}_k^2$. If $\overline{v_k}(t_0, x_0, p_0) \leq \mathcal{H}\overline{v_{k+1}}(t_0, x_0, p_0)$, then the subsolution inequality holds trivially. Now, if $\overline{v_k}(t_0, x_0, p_0) > \mathcal{H}\overline{v_{k+1}}(t_0, x_0, p_0)$, we argue by contradiction by assuming on the contrary that

$$\eta \;:=\; -\frac{\partial \varphi}{\partial t}(t_0, x_0, p_0) - \mathcal{L}\varphi(t_0, x_0, p_0) - f(x_0) \;>\; 0.$$

We set $p_0 = (t_i^0, e_i^0)_{1 \leq i \leq k}$. By continuity of φ and its derivatives, there exists some $\delta > 0$ with $t_0 + \delta < (t_1^0 + mh) \wedge T$ such that :

$$-\frac{\partial \varphi}{\partial t} - \mathcal{L}\varphi - f \;>\; \frac{\eta}{2}, \quad \text{on } ((t_0 - \delta, t_0 + \delta) \times B(x_0, \delta) \times B(p_0, \delta)) \cap \mathcal{D}_k^{2,m}. \quad (6.5)$$

From the definition of $\overline{v_k}$, there exists a sequence $(t_n, x_n, p_n)_{n\geq 1} \in ((t_0-\delta, t_0+\delta)\times B(x_0,\delta)\times B(p_0,\delta)) \cap \mathcal{D}_k^2$ such that $(t_n, x_n, p_n) \to (t_0, x_0, p_0)$ and $v_k(t_n, x_n, p_n) \to \overline{v_k}(t_0, x_0, p_0)$ as $n \to \infty$. By continuity of φ we also have that $\gamma_n := v_k(t_n, x_n, p_n) - \varphi(t_n, x_n, p_n)$ converges to 0 as $n \to \infty$. We set $p_n = (t_i^n, e_i^n)_{1\leq i\leq k}$. From the dynamic programming principle (3.8), for each $n \geq 1$, there exists a control $(\tau^n, \xi^n) \in \mathcal{I}_{t_n}$ such that

$$\begin{aligned} v_k(t_n, x_n, p_n) - \frac{\eta}{4}\delta_n \ &\leq \ \mathbb{E}\Big[\int_{t_n}^{\theta_n} f(X_s^n)ds + v_k(\theta_n, X_{\theta_n}^n, p_n)1_{\theta_n<\tau_n} \\ &\qquad + v_{k+1}(\theta_n, X_{\theta_n}, p_n\cup(\tau_n,\xi_n))1_{\tau_n\leq\theta_n}\Big]. \end{aligned} \tag{6.6}$$

Here $X^n := X^{t_n,x_n,0}$, we choose $\theta_n = \vartheta_n \wedge (t_n+\delta_n)$, with $\vartheta_n = \inf\{s \geq t_n : X_s^n \notin B(x_n, \frac{\delta}{2})\}$, and $(\delta_n)_n$ is a strictly positive sequence such that

$$\delta_n \to 0, \qquad \frac{\gamma_n}{\delta_n} \to 0, \quad \text{as } n \to \infty.$$

On the other hand, from Lemma 6.1, we have

$$\overline{\mathcal{H}v_{k+1}}(t_0, x_0, p_0) \ \leq \ \mathcal{H}\overline{v_{k+1}}(t_0, x_0, p_0) \ < \ \overline{v_k}(t_0, x_0, p_0) \ \leq \ \varphi(t_0, x_0, p_0).$$

Hence, since $\overline{\mathcal{H}v_{k+1}}$ is u.s.c. and φ is continuous, the inequality $\mathcal{H}v_{k+1} \leq \varphi$ holds in a neighborhood of (t_0, x_0, p_0), and so for sufficiently large n, we get :

$$v_{k+1}(\theta_n, X_{\theta_n}^n, p_n\cup(\tau_n,\xi_n))1_{\tau_n\leq\theta_n} \ \leq \ \varphi(\theta_n, X_{\theta_n}^n, p_n)1_{\tau_n\leq\theta_n} \quad a.s.$$

Together with (6.6), this yields :

$$\varphi(t_n, x_n, p_n) \ + \gamma_n - \frac{\eta}{4}\delta_n \ \leq \ \mathbb{E}\Big[\int_{t_n}^{\theta_n} f(X_s^n)ds + \varphi(\theta_n, X_{\theta_n}^n, p_n)\Big].$$

By applying Itô's formula to $\varphi(s, X_s^n, p_n)$ between $s = t_n$ and $s = \theta_n$, and dividing by δ_n, we then get :

$$\frac{\gamma_n}{\delta_n} - \frac{\eta}{4} \ \leq \ \frac{1}{\delta_n}\mathbb{E}\Big[\int_{t_n}^{\theta_n}\Big(\frac{\partial\varphi}{\partial t} + \mathcal{L}\varphi + f\Big)(s, X_s^n, p_n)ds\Big] \ \leq \ -\frac{\eta}{2}\mathbb{E}\Big[\frac{\theta_n - t_n}{\delta_n}\Big], \tag{6.7}$$

from (6.5). Now, from the growth linear condition on b, σ, Burkholder-Davis-Gundy inequality and Gronwall's lemma, we have the standard estimate : $\mathbb{E}[\sup_{s\in[t_n,t_n+\delta_n]}|X_s^n - x_n|^2] \to 0$, so that by Chebichev inequality, $\mathbb{P}[\vartheta_n \leq t_n+\delta_n] \to 0$, as n goes to infinity, and therefore by definition of θ_n :

$$1 \ \geq \ \mathbb{E}\Big[\frac{\theta_n - t_n}{\delta_n}\Big] \ \geq \ \mathbb{P}[\vartheta_n > t_n + \delta_n] \ \to \ 1, \quad \text{as } n \to \infty.$$

By sending n to infinity into (6.7), we obtain the required contradiction : $-\frac{\eta}{4} \leq -\frac{\eta}{2}$. □

6.3 Sequential comparison results

In this paragraph, we prove sequential comparison results. It involves some ideas of the iterative algorithm used to computed the value function numerically in the next chapter. First, we have to introduce some sets. For $k = 1, \ldots, m$, and any $n \geq 1$, we denote :

$$\begin{aligned}
\Theta_k(n) &= \Big\{t^{(k)} = (t_i)_{1 \leq i \leq k} \in \Theta_k \ : t_1 > T - nh\Big\}, \\
N &= \inf\{n \geq 1 \ : T - nh < 0\},
\end{aligned}$$

so that $\Theta_k(n)$ is strictly included in $\Theta_k(n+1)$ for $n = 1, \ldots, N-1$, and $\Theta_k(N) = \Theta_k$. We also denote for $k = 0$, and $n \geq 1$, $\mathbb{T}^n(0) = (T - nh, T] \cap [0, T]$ so that $\mathbb{T}^n(0) = (T - nh, T]$ is increasing with $n = 1, \ldots, N-1$, and $\mathbb{T}^N(0) = [0, T]$. We assumed $T - mh \geq 0$ to avoid trivialities so that $N > m$. We denote for $k = 0, \ldots, (n-m) \wedge m$, and $n = m, \ldots, N$,

$$\begin{aligned}
\mathcal{D}_k(n) &= \Big\{(t, x, p) \in \mathcal{D}_k \ : p \in \Theta_k(n) \times E^k\Big\}, \\
\mathcal{D}_k^i(n) &= \mathcal{D}_k(n) \cap \mathcal{D}_k^i \ = \ \Big\{(t, x, p) \in \mathcal{D}_k(n) \ : t \in \mathbb{T}_p^i(k)\Big\}, \ \ i = 1, 2,
\end{aligned}$$

with the convention that $\mathcal{D}_0(n) = \mathbb{T}^n(0) \times \mathbb{R}^d$, so that $\mathcal{D}_k(n)$ is strictly included in $\mathcal{D}_k(n+1)$ for $n = 1, \ldots, N-1$, and $\mathcal{D}_k(N) = \mathcal{D}_k$. We define sequential viscosity solutions as follows.

Definition 6.1. *Let $n \in \{m+1, \ldots, N\}$. We say that a family of locally bounded functions w_k on $\mathcal{D}_k(n)$, $k = 0, \ldots, m(n)$, is a viscosity supersolution (resp. subsolution) of* (4.1)-(4.2) *at step n if :*
(i) for all $k = 0, \ldots, m(n)$, $(t_0, x_0, p_0) \in \mathcal{D}_k^1(n)$, and $\varphi \in C^{1,2}(\mathcal{D}_k^1(n))$, which realizes a local minimum of $\underline{w_k} - \varphi$ (resp. maximum of $\overline{w_k} - \varphi$), we have

$$-\frac{\partial \varphi}{\partial t}(t_0, x_0, p_0) - \mathcal{L}\varphi(t_0, x_0, p_0) - f(x_0) \ \ \geq \ \ 0 \ \ (resp. \ \leq \ 0).$$

(ii) for all $k = 0, \ldots, m(n) - 1$, $(t_0, x_0, p_0) \in \mathcal{D}_k^2(n)$, and $\varphi \in C^{1,2}(\mathcal{D}_k^2(n))$, which realizes a local minimum of $\underline{w_k} - \varphi$ (resp. maximum of $\overline{w_k} - \varphi$), we have

$$\begin{aligned}
\min\Big\{ &-\frac{\partial \varphi}{\partial t}(t_0, x_0, p_0) - \mathcal{L}\varphi(t_0, x_0, p_0) - f(x_0) \ , \\
&\underline{w_k}(t_0, x_0, p_0) - \sup_{e \in E} \underline{w_{k+1}}(t_0, x_0, p_0 \cup (t_0, e))\Big\} \ \ \geq \ \ 0
\end{aligned}$$

(resp.

$$\begin{aligned}
\min\Big\{ &-\frac{\partial \varphi}{\partial t}(t_0, x_0, p_0) - \mathcal{L}\varphi(t_0, x_0, p_0) - f(x_0) \ , \\
&\overline{w_k}(t_0, x_0, p_0) - \sup_{e \in E} \overline{w_{k+1}}(t_0, x_0, p_0 \cup (t_0, e))\Big\} \ \ \leq \ \ 0).
\end{aligned}$$

We say that a family of locally bounded functions w_k on $\mathcal{D}_k(n)$, $k = 0, \ldots, m(n)$, is a viscosity solution of (4.1)-(4.2) *at step n if it is a viscosity supersolution and subsolution of* (4.1)-(4.2) *at step n.*

We then prove the following comparison principle at step n.

Proposition 6.3. *Let $n \in \{m+1, \ldots, N\}$. Let u_k (resp. w_k), $k = 0, \ldots, m(n)$, be a family of viscosity subsolution (resp. supersolution) of* (4.1)-(4.2) *at step n satisfying growth condition* (2.9). *Suppose also that w_k satisfies* (4.3). *If u_k and w_k are such that for all $x \in \mathbb{R}^d$*

$$\begin{aligned} \overline{u_k}(t_1+mh, x, p) &\leq \underline{w_k}(t_1+mh, x, p), \quad p = (t_i, e_i)_{1\leq i\leq k} \in \Theta_k(n) \times E^k, \ k \geq 1, \\ \overline{u_0}(T, x) &\leq \underline{w_0}(T, x). \end{aligned}$$

Then, $\overline{u_k} \leq \underline{w_k}$ on $\mathcal{D}_k(n)$, $k = 0, \ldots, m(n)$.

Remark 6.1. We recall some basic definitions and properties in viscosity solutions theory, which shall be used in the proof of the above proposition. Consider the general PDE

$$F(t, x, w, \frac{\partial w}{\partial t}, D_x w, D_x^2 w) \;=\; 0 \quad \text{on } [t_0, t_1) \times \mathcal{O}, \tag{6.8}$$

where $t_0 < t_1$, and $\mathcal{O}$ is an open set in $\mathbb{R}^d$. There is an equivalent definition of viscosity solutions to (6.8) in terms of semi-jets $\bar{J}^{2,+}w(t,x)$ and $\bar{J}^{2,-}w(t,x)$ associated respectively to an upper-semicontinuous (u.s.c.) and lower-semicontinuous (l.s.c.) function w (see [27] or [38] for the definition of semi-jets) : an u.s.c. (resp. l.s.c.) function w is a viscosity subsolution (resp. supersolution) to (6.8) if and only if for all $(t,x) \in [t_0, t_1) \times \mathcal{O}$,

$$F(t, x, w(t,x), r, q, A) \;\leq\; (\text{ resp. } \geq) \;\; 0, \quad \forall (r, q, A) \in \bar{J}^{2,+}w(t,x) \; (\text{ resp. } \bar{J}^{2,-}w(t,x)).$$

For $\eta > 0$, we say that w^η is a viscosity η-strict supersolution to (6.8), if w^η is a viscosity supersolution to

$$F(t, x, w^\eta, \frac{\partial w^\eta}{\partial t}, D_x w^\eta, D_x^2 w^\eta) \;\geq\; \eta, \quad \text{on } [t_0, t_1) \times \mathcal{O}.$$

in the sense that it is a viscosity supersolution to $F(t, x, w^\eta, \frac{\partial w^\eta}{\partial t}, D_x w^\eta, D_x^2 w^\eta) - \eta = 0$, on $[t_0, t_1) \times \mathcal{O}$.

As usual when dealing with variational inequalities, we begin the proof of the comparison principle by showing the existence of viscosity η-strict supersolutions for equation (4.1)-(4.2).

Lemma 6.2. *Let $n \in \{m+1, \ldots, N\}$. Let w_k, $k = 0, \ldots, m(n)$, be a family of viscosity supersolutions of* (4.1)-(4.2) *satisfying* (4.3). *Then, for any $\eta > 0$, there exists a family of viscosity η-strict supersolutions w_k^η of* (4.1)-(4.2) *such that for $k = 0, \ldots, m(n)$:*

$$w_k(t,x,p) + \eta C_1|x|^2 + \eta h_{k,n}(t,t_1) \leq w_k^\eta(t,x,p) \leq w_k(t,x,p) + \eta C_2(1+|x|^2) + \eta h_{k,n}(t,t_1) \tag{6.9}$$

for all $(t,x,p) \in \mathcal{D}_k$, for some positive constants C_1, C_2 independent on η, with

$$h_{k,n}(t,t_1) = 1_{k\geq 1}\frac{1}{t_1 - T + nh} + 1_{k=0}\frac{1}{t - T + nh}.$$

Moreover, for $k = 0, \ldots, m(n)-1$, $(t,x,p) \in \mathcal{D}_k(n)$, $p = (t_i, e_i)_{1\leq i\leq k}$ with $t = t_k + h$, we have :

$$\underline{w_k^\eta}(t,x,p) \;\geq\; \sup_{e\in E} \underline{w_{k+1}^\eta}(t, x, p \cup (t,e)) + \eta. \tag{6.10}$$

Proof. For $\eta > 0$, consider the functions :

$$\begin{aligned} w_k^\eta(t,x,p) &= w_k(t,x,p) + \eta\phi_{1,k}(t) + \eta\phi_2(t,x) + \eta\phi_{3,k}(t,t_1), \\ \phi_{1,k}(t) &= \left[(T-t) + (m-k)\right], \\ \phi_2(t,x) &= \frac{1}{2}e^{L(T-t)}\left(1+|x|^2\right), \\ \phi_{3,k}(t,t_1) &= 1_{k\geq 1}\frac{1}{t_1 - T + nh} + 1_{k=0}\frac{1}{t-T+nh} \end{aligned}$$

with L a positive constant to be determined later. It is clear that w_k^η satisfies (6.9) with $C_1 = 1/2$ and $C_2 = T + m + e^{LT}/2$. Moreover, we easily show that $w_k + \eta\phi_{1,k}^\eta$ is a viscosity supersolution to

$$-\frac{\partial(w_k + \eta\phi_{1,k})}{\partial t} - \mathcal{L}(w_k + \eta\phi_{1,k}) - f \quad \geq \quad \eta. \tag{6.11}$$

This is derived from the fact that $-\frac{\partial \phi_{1,k}}{\partial t} - \mathcal{L}\phi_{1,k} = 1$, and w_k is a viscosity supersolution to $-\frac{\partial w_k}{\partial t} - \mathcal{L}w_k - f \geq 0$. We now show that ϕ_2 is a supersolution to

$$-\frac{\partial \phi_2}{\partial t} - \mathcal{L}\phi_2 \quad \geq \quad 0. \tag{6.12}$$

This is done by calculating this quantity explicitely. Indeed, we have

$$\frac{\partial \phi_2}{\partial t}(t,x) = -\frac{L}{2}e^{L(T-t)}(1+|x|^2), \quad \mathcal{L}\phi_2(t,x) = e^{L(T-t)}\left(b(x).x + \mathrm{tr}\left(\sigma\sigma'(x)\right)\right).$$

Since b and σ are of linear growth, we thus obtain :

$$-\frac{\partial \phi_2}{\partial t}(t,x) - \mathcal{L}\phi_2(t,x) \quad \geq \quad e^{L(T-t)}\left[\frac{L}{2}(1+|x|^2) - C(1+|x|+|x|^2)\right],$$

for some constant C independent of t, x. Therefore, by taking L sufficiently large, we get the required inequality (6.12). furthermore we have:

$$\frac{\partial \phi_3}{\partial t}(t,t_1) \leq 0$$

which shows together with (6.11) that w_k^η is a viscosity supersolution to

$$-\frac{\partial w_k^\eta}{\partial t} - \mathcal{L}w_k^\eta - f \quad \geq \quad \eta. \tag{6.13}$$

Moreover, since

$$\underline{w_k}(t,x,p) - \sup_{e\in E}\underline{w_{k+1}}(t,x,p\cup(t,e)) \quad \geq \quad 0,$$

we immediately get

$$\begin{aligned} & \underline{w_k^\eta}(t,x,p) - \sup_{e\in E}\underline{w_{k+1}^\eta}(t,x,p\cup(t,e)) \\ = \quad & \underline{w_k}(t,x,p) + \eta\phi_{1,k}(t) - \sup_{e\in E}\underline{w_{k+1}}(t,x,p\cup(t,e)) - \eta\phi_{1,k+1}(t) \\ \geq \quad & \eta\phi_{1,k}(t) - \eta\phi_{1,k+1}(t) \ \geq \ \eta. \end{aligned}$$

Together with (6.13), this proves the required viscosity η-strict supersolution property for w_k^η to (4.1)-(4.2). □

The main step in the proof of Proposition 6.3 consists in the comparison principle for η-strict supersolutions. Notice from (6.9) that once w_k satisfies a linear growth condition, then w_k^η satisfies the quadratic growth lower-bound condition :

$$\eta C_1 |x|^2 - C_2 + \frac{\eta}{t_1 - T + nh} \quad \leq \quad w_k^\eta(t,x,p), \quad (t,x,p) \in \mathcal{D}_k, \tag{6.14}$$

for some positive constants C_1, C_2.

Lemma 6.3. *Let $n \in \{m+1, \ldots, N\}$ and $\eta > 0$. Let u_k (resp. w_k), $k = 0, \ldots, (n-m) \wedge m$, be a family of viscosity subsolution (resp. η-strict supersolution) of* (4.1)-(4.2) *at step n, with u_k satisfying the linear growth condition* (2.9) *and w_k satisfying the quadratic growth condition* (6.14)*. Suppose that for all $x \in \mathbb{R}^d$,*

$$\begin{aligned}
\overline{u_k}(t_1 + mh, x, p) &\leq \underline{w_k}(t_1 + mh, x, p), \quad p = (t_i, e_i)_{1 \leq i \leq k} \in \Theta_k(n) \times E^k, k \geq 1, & (6.15)\\
\overline{u_0}(T, x) &\leq \underline{w_0}(T, x). & (6.16)\\
\underline{w}_k(t_k + h, x, \pi) &\geq \sup_{e \in E} \underline{w_{k+1}}\,(t_k + h, x, p \cup (t_k + h, e)) + \eta, & (6.17)\\
&\quad p = (t_i, e_i)_{1 \leq i \leq k} \in \Theta_k(n) \times E^k, \ k \leq m - 1. &\\
\underline{w}_k(T - mh, x, \pi) &\geq \sup_{e \in E} \underline{w_{k+1}}\,(T - mh, x, p \cup (T - mh, e)) + \eta, & (6.18)\\
&\quad \textit{for all } (T - mh, x, p) \in \mathcal{D}_k(n). & (6.19)
\end{aligned}$$

Then, $\overline{u_k} \leq \underline{w_k}$ on $\mathcal{D}_k(n)$, $k = 0, \ldots, (n-m) \wedge m$.

Proof. From the linear growth of u_k, and from the quadratic growth lower-bound of w_k, we have

$$\overline{u_k}(t,x,p) - \underline{w_k}(t,x,p) \quad \leq \quad C_1 \left(1 + |x|\right) - C_2 \, |x|^2 - \frac{\eta}{1_{k \geq 1} t_1 + 1_{k=0} t - T + nh},$$

for all $k = 0, \ldots, m$, $(t,x,p) \in \mathcal{D}_k(n)$, for some positive constants C_1, C_2. Thus, for all k, the supremum of the u.s.c function $\overline{u_k} - \underline{w_k}$ is attained on a compact set that only depends on C_1 and C_2. Hence, one can find $k_0 \in \{0, \ldots, (n-m) \wedge m\}$, $(t_0, x_0, p_0) \in \mathcal{D}_{k_0}(n)$ such that :

$$\begin{aligned}
M \quad &:= \sup_{\substack{k \,\in\, \{0, \ldots, m\} \\ (t,x,p) \,\in\, \mathcal{D}_k(n)}} \left[\overline{u_k}(t,x,p) - \underline{w_k}(t,x,p)\right] \\
&= \quad \overline{u_{k_0}}(t_0, x_0, p_0) - \underline{w_{k_0}}(t_0, x_0, p_0), \qquad (6.20)
\end{aligned}$$

and we have to show that $M \leq 0$. We set $p_0 = (t_i^0, e_i^0)_{1 \leq i \leq k_0}$, and we distinguish the six possible cases concerning (k_0, t_0, x_0, p_0) :

- *Case 1 : $k_0 \neq 0$, $t_0 = t_1^0 + mh$.*
- *Case 2 : $k_0 = 0$, $t_0 = T$.*
- *Case 3 : $k_0 \neq 0$, $t_0 \in \mathbb{T}^1_{p_0}(k_0)$.*

- *Case 4 :* $k_0 = 0$, $t_0 \in [0, T - mh)$ *or* $k_0 \in \{1, \ldots, m-1\}$, $t_0 \in \mathbb{T}^2_{p_0}(k_0)$, $t_0 \neq t^0_{k_0} + h$, $t_0 \neq T - mh$.
- *Case 5 :* $k_0 \in \{1, \ldots, m-1\}$, $t_0 = t^0_{k_0} + h$.
- *Case 6 :* $k_0 \in \{1, \ldots, m-1\}$, $t_0 = T - mh$, .

▶ *Cases 1 and 2 :* these two cases imply directly from (6.15) (resp. (6.16)) that $M \leq 0$.

▶ *Cases 3 and 4 :* we focus only on case 4, as case 3 involves similar (and simpler) arguments. We follow general viscosity solution technique based on the Ishii technique and work towards a contradiction. To this end, let us consider the following function :

$$\Phi_\varepsilon(t, t', x, x', p, p') = \overline{u_{k_0}}(t, x, p) - \underline{w_{k_0}}(t', x', p') - \psi_\varepsilon(t, t', x, x', p, p'),$$

with

$$\begin{aligned}\psi_\varepsilon(t, t', x, x', p, p') &= \frac{1}{2}\big[|t - t_0|^2 + |p - p_0|^2\big] + \frac{1}{4}|x - x_0|^4 \\ &\quad + \frac{1}{2\varepsilon}\big[[t - t'|^2 + |x - x'|^2 + [p - p'|^2\big].\end{aligned}$$

By the positiveness of the function ψ_ε, we notice that (t_0, x_0, p_0) is a strict maximizer of $(t, x, p) \to \Phi_\varepsilon(t, t, x, x, p, p)$. Hence, by Proposition 3.7 in [27], there exists a sequence of maximizers $(t_\varepsilon, t'_\varepsilon, x_\varepsilon, x'_\varepsilon, p_\varepsilon, p'_\varepsilon)$ of Φ_ε such that :

$$\begin{aligned}(t_\varepsilon, t'_\varepsilon, x_\varepsilon, x'_\varepsilon, p_\varepsilon, p'_\varepsilon) &\to (t_0, t_0, x_0, x_0, p_0, p_0), && (6.21)\\ \overline{u_{k_0}}(t_\varepsilon, x_\varepsilon, p_\varepsilon) - \underline{w_{k_0}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon) &\to \overline{u_{k_0}}(t_0, x_0, p_0) - \underline{w_{k_0}}(t_0, x_0, p_0), && (6.22)\\ \frac{1}{\varepsilon}\big[|t_\varepsilon - t'_\varepsilon|^2 + |x_\varepsilon - x'_\varepsilon|^2 + |p_\varepsilon - p'_\varepsilon|^2\big] &\to 0 \text{ as } \varepsilon \to 0. && (6.23)\end{aligned}$$

By applying Theorem 3.2 in [27] to the sequence of maximizers $(t_\varepsilon, t'_\varepsilon, x_\varepsilon, x'_\varepsilon, p_\varepsilon, p'_\varepsilon)$ of Φ_ε, we get the existence of two symmetric matrices $A_\varepsilon, A'_\varepsilon$ such that :

$$\begin{aligned}(r_\varepsilon, q_\varepsilon, A_\varepsilon) &\in \overline{J}^{2,+}\overline{u_{k_0}}(t_\varepsilon, x_\varepsilon, p_\varepsilon) && (6.24)\\ \left(r'_\varepsilon, q'_\varepsilon, A'_\varepsilon\right) &\in \overline{J}^{2,-}\underline{w_{k_0}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon), && (6.25)\end{aligned}$$

where

$$\begin{aligned}r_\varepsilon &= \frac{\partial \psi_\varepsilon}{\partial t}(t_\varepsilon, t'_\varepsilon, x_\varepsilon, x'_\varepsilon, p_\varepsilon, p'_\varepsilon) = \frac{1}{\varepsilon}(t_\varepsilon - t'_\varepsilon) + (t_\varepsilon - t_0), && (6.26)\\ r'_\varepsilon &= -\frac{\partial \psi_\varepsilon}{\partial t'}(t_\varepsilon, t'_\varepsilon, x_\varepsilon, x'_\varepsilon, p_\varepsilon, p'_\varepsilon) = \frac{1}{\varepsilon}(t_\varepsilon - t'_\varepsilon) && (6.27)\\ q_\varepsilon &= \frac{\partial \psi_\varepsilon}{\partial x}(t_\varepsilon, t'_\varepsilon, x_\varepsilon, x'_\varepsilon, p_\varepsilon, p'_\varepsilon) = \frac{1}{\varepsilon}(x_\varepsilon - x'_\varepsilon) + |x_\varepsilon - x_0|^2\,(x_\varepsilon - x_0), && (6.28)\\ q'_\varepsilon &= -\frac{\partial \psi_\varepsilon}{\partial x'}(t_\varepsilon, t'_\varepsilon, x_\varepsilon, x'_\varepsilon, p_\varepsilon, p'_\varepsilon) = \frac{1}{\varepsilon}(x_\varepsilon - x'_\varepsilon), && (6.29)\end{aligned}$$

and

$$\begin{pmatrix} A_\varepsilon & 0 \\ 0 & -A'_\varepsilon \end{pmatrix} \leq \begin{pmatrix} \frac{3}{\varepsilon}I_d - Q\,(x_\varepsilon - x_0) & -\frac{3}{\varepsilon}I_d \\ -\frac{3}{\varepsilon}I_d & \frac{3}{\varepsilon}I_d \end{pmatrix}, \quad (6.30)$$

with

$$Q(x) \quad = \quad 2x \otimes x + |x|^2 I_d,$$

I_d the identity matrix of dimension $d \times d$, and for $x = (x_i)_{1 \leq i \leq d} \in \mathbb{R}^d$, $x \otimes x$ is the tensorial product defined by $x \otimes x = (x_i x_j)_{i,j \in \{1..d\}^2}$. Here, to alleviate notations, and since there is no derivatives with respect to the variable p in the PDE, the semi-jets are defined with respect to the variables (t,x), and we omitted the terms corresponding to the derivatives of ψ_ε with respect to p. We set $p_\varepsilon = (t_i^\varepsilon, e_i^\varepsilon)_{1 \leq i \leq k_0}$, and $p'_\varepsilon = (t_i^{\prime\varepsilon}, e_i^{\prime\varepsilon})_{1 \leq i \leq k_0}$. From (6.21), we deduce that for ε small enough, $t_\varepsilon \in \mathbb{T}^2_{p_0}(k_0)$ and $t_\varepsilon \neq t^\varepsilon_{k_0} + h$. From (6.24)-(6.25), and the formulation of viscosity subsolution of u_{k_0} to (4.2) and η-strict viscosity supersolution of w_{k_0} to (4.2) by means of semi-jets, we have for all ε small enough :

$$\min\Big\{ -r_\varepsilon - b(x_\varepsilon) q_\varepsilon - \frac{1}{2} \mathrm{tr}\left(\sigma\sigma'(x_\varepsilon) A_\varepsilon\right) - f(x_\varepsilon), \quad \overline{u_{k_0}}(t_\varepsilon, x_\varepsilon, p_\varepsilon) - \sup_{e \in E} \overline{u_{k_0+1}}(t_\varepsilon, x_\varepsilon, p_\varepsilon \cup (t_\varepsilon, e)) \Big\} \quad \leq \quad 0, \tag{6.31}$$

$$\min\Big\{ -r'_\varepsilon - b(x'_\varepsilon) q'_\varepsilon - \frac{1}{2} \mathrm{tr}\left(\sigma\sigma'(x'_\varepsilon) A'_\varepsilon\right) - f(x'_\varepsilon), \quad \underline{w_{k_0}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon) - \sup_{e \in E} \underline{w_{k_0+1}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon \cup (t'_\varepsilon, e)) \Big\} \quad \geq \quad \eta. \tag{6.32}$$

We then distinguish the following two possibilities in (6.31) :

- (i) for all ε small enough,

  $$\overline{u_{k_0}}(t_\varepsilon, x_\varepsilon, p_\varepsilon) - \sup_{e \in E} \overline{u_{k_0+1}}(t_\varepsilon, x_\varepsilon, p_\varepsilon \cup (t_\varepsilon, e)) \quad \leq \quad 0.$$

  Then, for all ε small enough, there exists $e_\varepsilon \in E$ such that :

  $$\overline{u_{k_0}}(t_\varepsilon, x_\varepsilon, p_\varepsilon) \quad \leq \quad \overline{u_{k_0+1}}(t_\varepsilon, x_\varepsilon, p_\varepsilon \cup (t_\varepsilon, e_\varepsilon)) + \frac{\eta}{2}.$$

  Moreover, by (6.32), we have

  $$\underline{w_{k_0}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon) \quad \geq \quad \underline{w_{k_0+1}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon \cup (t'_\varepsilon, e_\varepsilon)) + \eta.$$

  Combining the two above inequalities, we deduce that for all ε small enough,

  $$\begin{aligned} & \overline{u_{k_0}}(t_\varepsilon, x_\varepsilon, p_\varepsilon) - \underline{w_{k_0}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon) \\ \leq \quad & \overline{u_{k_0+1}}(t_\varepsilon, x_\varepsilon, p_\varepsilon \cup (t_\varepsilon, e_\varepsilon)) - \underline{w_{k_0+1}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon \cup (t'_\varepsilon, e_\varepsilon)) - \frac{\eta}{2}. \end{aligned}$$

  Since E is compact, there exists some $e \in E$ s.t. $e_\varepsilon \to e$ up to a subsequence. From (6.21)-(6.22), and since $\overline{u_{k_0}}$, $-\underline{w_{k_0}}$ are u.s.c., we obtain by sending ε to zero :

  $$\begin{aligned} & \overline{u_{k_0}}(t_0, x_0, p_0) - \underline{w_{k_0}}(t_0, x_0, p_0) \\ \leq \quad & \overline{u_{k_0+1}}(t_0, x_0, p_0 \cup (t_0, e)) - \underline{w_{k_0+1}}(t_0, x_0, p_0 \cup (t_0, e)) - \frac{\eta}{2}, \end{aligned}$$

  which contradicts (6.20).

- (ii) for all ε small enough,

$$-r_\varepsilon - b(x_\varepsilon)q_\varepsilon - \frac{1}{2}\mathrm{tr}\left(\sigma\sigma'(x_\varepsilon)A_\varepsilon\right) - f(x_\varepsilon) \quad \leq \quad 0.$$

Combining with (6.32), we then get

$$\begin{aligned} \eta \quad \leq \quad & r_\varepsilon - r'_\varepsilon + b(x_\varepsilon)q_\varepsilon - b(x'_\varepsilon)q'_\varepsilon \\ & + \frac{1}{2}\mathrm{tr}\left(\sigma\sigma'(x_\varepsilon)A_\varepsilon - \sigma\sigma'(x'_\varepsilon)A'_\varepsilon\right) + f(x_\varepsilon) - f(x'_\varepsilon). \end{aligned} \tag{6.33}$$

We now analyze the convergence of the r.h.s. of (6.33) as ε goes to zero. First, we see from (6.21) and (6.26)-(6.27) that $r_\varepsilon - r'_\varepsilon$ converge to zero. We also immediately see from the continuity of f and (6.21) that $f(x_\varepsilon) - f(x'_\varepsilon)$ converge to zero. It is also clear from the Lipschitz property of b, (6.21), (6.23), and (6.28)-(6.29) that $b(x_\varepsilon)q_\varepsilon - b(x'_\varepsilon)q'_\varepsilon$ converge to zero. Finally, from (6.30), we have

$$\begin{aligned} \mathrm{tr}\left(\sigma\sigma'(x_\varepsilon)A_\varepsilon - \sigma\sigma'(x'_\varepsilon)A'_\varepsilon\right) \quad \leq \quad & \frac{3}{\varepsilon}\mathrm{tr}\left((\sigma(x_\varepsilon) - \sigma(x'_\varepsilon))(\sigma(x_\varepsilon) - \sigma(x'_\varepsilon))'\right) \\ & - \mathrm{tr}\left(\sigma\sigma'(x_\varepsilon)Q(x_\varepsilon - x_0)\right), \end{aligned}$$

and the r.h.s. of the above inequality converges to zero from the Lipschitz property of σ, (6.21) and (6.23). Therefore, by sending ε to zero into (6.33), we obtain the required contradiction : $\eta \leq 0$.

► *Case 5 and 6* : We only consider the proof of case 5, as case 6 is similar. We keep the same notations as in the previous case. The crucial difference is that $\overline{u_{k_0}}$ and $\underline{w_{k_0}}$ may be sub and supersolution to different equations, depending on the position of t_ε (resp. t'_ε) with respect to $t^\varepsilon_{k_0} + h$ (resp. $t'^\varepsilon_{k_0} + h$). Actually, up to a subsequence for ε, we have three subcases. If $t_\varepsilon \geq t^\varepsilon_{k_0} + h$ and $t'_\varepsilon \geq t'^\varepsilon_{k_0} + h$ for all ε small enough, the proof of the preceding case applies. If $t_\varepsilon < t^\varepsilon_{k_0} + h$, for all ε small enough, then we have the viscosity subsolution (resp. supersolution) property of $\overline{u_{k_0}}$ (resp. $\underline{w_{k_0}}$) to the same linear PDE : $-\dfrac{\partial v_k}{\partial t} - \mathcal{L}v_k - f = 0$, at $(t_\varepsilon, x_\varepsilon, p_\varepsilon)$ (resp. $(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon)$), and we conclude as in *Case 3*. Finally, if $t_\varepsilon \geq t^\varepsilon_{k_0} + h$ and $t'_\varepsilon < t'^\varepsilon_{k_0} + h$ for all ε small enough, then the viscosity subsolution property of u_{k_0} to (4.2) at $(t_\varepsilon, x_\varepsilon, p_\varepsilon)$, and the viscosity η-strict supersolution property of w_{k_0} to (4.1) at $(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon)$ lead to :

$$-r'_\varepsilon - b(x'_\varepsilon)q'_\varepsilon - \frac{1}{2}\mathrm{tr}\left(\sigma\sigma'(x'_\varepsilon)A'_\varepsilon\right) - f(x'_\varepsilon) \quad \geq \quad \eta \tag{6.34}$$

and the following two possibilities :

$$-r_\varepsilon - b(x_\varepsilon)q_\varepsilon - \frac{1}{2}\mathrm{tr}\left(\sigma\sigma'(x_\varepsilon)A_\varepsilon\right) - f(x_\varepsilon) \quad \leq \quad 0, \tag{6.35}$$

or

$$\overline{u_{k_0}}(t_\varepsilon, x_\varepsilon, p_\varepsilon) - \sup_{e \in E} \overline{u_{k_0+1}}(t_\varepsilon, x_\varepsilon, p_\varepsilon \cup (t_\varepsilon, e)) \quad \leq \quad 0. \tag{6.36}$$

The first possibility (6.34), (6.35) is dealt with by the same arguments as in *Case 4* (ii). The second possibility (6.34), (6.36) does not allow to conclude directly. In fact, we use the additional condition (6.17) :

$$\underline{w_{k_0}}(t_0, x_0, p_0) \quad \geq \quad \sup_{e\in E} \underline{w_{k_0+1}}(t_0, x_0, p_0 \cup (t_0, e)) + \eta. \tag{6.37}$$

Since $\underline{w_{k_0}}$ is lower semicontinuous, this implies by (6.21) that for all ε small enough :

$$\begin{aligned}\underline{w_{k_0}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon) \quad &\geq \quad \underline{w_{k_0}}(t_0, x_0, p_0) - \frac{\eta}{2} \\ &\geq \quad \sup_{e\in E} \underline{w_{k_0+1}}(t_0, x_0, p_0 \cup (t_0, e)) + \frac{\eta}{2}.\end{aligned}$$

Hence, by combining with (6.36), we deduce that

$$\begin{aligned}&\overline{u_{k_0}}(t_\varepsilon, x_\varepsilon, p_\varepsilon) - \underline{w_{k_0}}(t'_\varepsilon, x'_\varepsilon, p'_\varepsilon) + \frac{\eta}{2} \\ \leq \quad &\sup_{e\in E} \overline{u_{k_0+1}}(t_\varepsilon, x_\varepsilon, p_\varepsilon \cup (t_\varepsilon, e)) - \sup_{e\in E} \underline{w_{k_0+1}}(t_0, x_0, p_0 \cup (t_0, e)),\end{aligned}$$

for all ε small enough. From (6.22) and Lemma 6.1, we then obtain by sending ε to zero :

$$\begin{aligned}&\overline{u_{k_0}}(t_0, x_0, p_0) - \underline{w_{k_0}}(t_0, x_0, p_0) + \frac{\eta}{2} \\ \leq \quad &\sup_{e\in E} \overline{u_{k_0+1}}(t_0, x_0, p_0 \cup (t_0, e)) - \sup_{e\in E} \underline{w_{k_0+1}}(t_0, x_0, p_0 \cup (t_0, e)) \\ \leq \quad &\sup_{e\in E} \left\{ \overline{u_{k_0+1}}(t_0, x_0, p_0 \cup (t_0, e)) - \underline{w_{k_0+1}}(t_0, x_0, p_0 \cup (t_0, e)) \right\}.\end{aligned}$$

This is in contradiction with (6.20). □

Finally, as usual, the comparison theorem for strict supersolutions implies comparison for supersolutions.

Proof of Proposition 6.3

For any $\eta > 0$, we use Lemma 6.2 to obtain an η-strict supersolution w_k^η of (4.1)-(4.2), which satisfies (6.9), so that $\underline{w_k}(t, x, p) \to \underline{w_k^\eta}(t, x, p)$ for all $(t, x, p) \in \mathcal{D}_k$, as η goes to zero. We then use Lemma 6.3 to deduce that $\overline{u_k} \leq \underline{w_k}^\eta$ on $\mathcal{D}_k(n)$, $k = 0, \ldots, (n-m) \wedge m$. Thus, letting $\eta \to 0$, completes the proof. □

6.4 Boundary data and continuity

In this paragraph, we shall derive by induction the boundary data (4.4)-(4.5) in Proposition 4.2, and the continuity of the value functions as byproducts of viscosity properties and sequential comparison principles.

We first show relation (4.5), which follows easily from the definition of the value functions.

Lemma 6.4. *For all $x \in \mathbb{R}^d$, $v_0(T^-, x)$ exists and is equal to:*

$$v_0(T^-, x) = g(x) \tag{6.38}$$

Proof. For any $(t,x) \in (T-mh,T)\times\mathbb{R}^d$, we have from the definition of v_0, and the fact that no order can be passed after $T-mh$:

$$v_0(t,x) = \mathbb{E}\Big[\int_t^T f(X_s^{t,x,0})ds + g(X_T^{t,x,0})\Big].$$

Therefore, with the continuity and linear growth assumptions on f and g, we get the result from the dominated convergence theorem. □

The derivation of relation (4.4) is more delicate. We first state the following result, which is a direct consequence of the dynamic programming principle.

Lemma 6.5. *(i) For $k=1,\ldots,m$, and $p=(t_i,e_i)_{1\le i\le k}\in\Theta_k\times E^k$, we have for all $x\in\mathbb{R}^d$, and $t\in[t_k,(t_k+h)\wedge(t_1+mh))$,*

$$\begin{aligned} v_k(t,x,p) &= \mathbb{E}\Big[\int_t^{(t_k+h)\wedge(t_1+mh)} f(X_s^{t,x,0})ds + v_k(t_k+h,X_{t_k+h}^{t,x,0},p)1_{t_k+h<t_1+mh} \qquad (6.39)\\ &\quad + \Big(c(X_{t_1+mh}^{t,x,0},e_1)+v_{k-1}(t_1+mh,\Gamma(X_{t_1+mh}^{t,x,0},e_1),p_-)\Big)1_{t_1+mh\le t_k+h}\Big]. \end{aligned}$$

(ii)For $k=1,\ldots,m$, and $p=(t_i,e_i)_{1\le i\le k}\in\Theta_k\times E^k$, we have for all $x\in\mathbb{R}^d$, and $t\in(T-mh,T)$,

$$\begin{aligned} v_k(t,x,p) &= \mathbb{E}\Big[\int_t^{t_1+mh} f(X_s^{t,x,0})ds + c(X_{t_1+mh}^{t,x,0},e_1) \qquad (6.40)\\ &+ v_{k-1}(t_1+mh,\Gamma(X_{t_1+mh}^{t,x,0},e_1),p_-)\Big]. \end{aligned}$$

(iii) For $k=1,\ldots,m$, and $p=(t_i,e_i)_{1\le i\le k}\in\Theta_k\times E^k$, such that $t_k+h<t_1+mh$ and $t_k+h\le T-mh$, we have for all $x\in\mathbb{R}^d$, and $t\in\mathbb{T}_p^2(k)=[t_k+h,t_1+mh)\cap[0,T-mh]$,

$$\begin{aligned} v_k(t,x,p) &\ge \mathbb{E}\Big[\int_t^{t_1+mh} f(X_s^{t,x,0})ds\\ &\qquad + c(X_{t_1+mh}^{t,x,0},e_1)+v_{k-1}(t_1+mh,\Gamma(X_{t_1+mh}^{t,x,0},e_1),p_-)\Big] \qquad (6.41)\\ v_k(t,x,p) &\le \sup_{(\tau,\xi)\in\mathcal{I}_t}\mathbb{E}\Big[\int_t^{(t_1+mh)\wedge\tau} f(X_s^{t,x,0})ds + v_{k+1}(\tau,X_\tau^{t,x,0},p\cup(\tau,\xi))1_{\tau<t_1+mh}\\ &\quad +\Big(c(X_{t_1+mh}^{t,x,0},e_1)+v_{k-1}(t_1+mh,\Gamma(X_{t_1+mh}^{t,x,0},e_1),p_-)\Big)1_{t_1+mh\le\tau}\Big]. \qquad (6.42) \end{aligned}$$

Proof. First, we recall from the dynamic programming principle that by making an immediate impulse control, i.e. by taking in (3.8), $\theta=t$ and $\tau=t$, $\xi=e$ arbitrary in E, we have for all $k=0,\ldots,m-1$, $p=(t_i,e_i)_{1\le i\le k}\in\Theta_k\times E^k$, $(t,x)\in\mathbb{T}_p(k)\times\mathbb{R}^d$ with $t\ge t_k+h$,

$$v_k(t,x,p) \ \ge\ \sup_{e\in E} v_{k+1}(t,x,p\cup(t,e)). \qquad (6.43)$$

(i) Fix $k=1,\ldots,m$, $p=(t_i,e_i)_{1\le i\le k}\in\Theta_k\times E^k$, and $(t,x)\in\mathbb{T}_p^1(k)\times\mathbb{R}^d$. We distinguish the two following cases :

• *Case 1* : $t_k + h < t_1 + mh$. Then, for all $\alpha \in \mathcal{A}_{t,p}$, we have from (2.3), $X_s^{t,x,p,\alpha} = X_s^{t,x,0}$ for $t \leq s \leq t_k + h$. Hence, by applying (3.4) with $\theta = t_k + h$, and noting that $\tau_i + mh > \theta$, $k(\theta,\alpha) = k$, $p(\theta,\alpha) = p$ for any $\alpha = (\tau_i,\xi_i) \in \mathcal{A}_{t,p}$, we obtain the required relation (6.39), i.e.

$$v_k(t,x,p) \;=\; \mathbb{E}\Big[\int_t^{t_k+h} f(X_s^{t,x,0})ds + v_k(t_k+h, X_{t_k+h}^{t,x,0}, p)\Big].$$

• *Case 2* : $t_1 + mh \leq t_k + h$. Then, for all $\alpha \in \mathcal{A}_{t,p}$, we have from (2.3), $X_s^{t,x,p,\alpha} = X_s^{t,x,0}$ for $t \leq s < t_1 + mh$, and $X_{t_1+mh}^{t,x,p,\alpha} = \Gamma(X_{t_1+mh}^{t,x,0}, e_1)$. Hence, by applying (3.4) with $\theta = t_1 + mh$, and noting that for any $\alpha = (\tau_i,\xi_i) \in \mathcal{A}_{t,p}$, we have either $k(\theta,\alpha) = k-1$, $p(\theta,\alpha) = p_-$ if $\tau_{k+1} > t_1 + mh$ (which always arises when $t_1 + mh < t_k + h$), or $k(\theta,\alpha) = k$, $p(\theta,\alpha) = p_- \cup (\tau_{k+1},\xi_{k+1})$ if $\tau_{k+1} = t_k + h = t_1 + mh$, we obtain

$$\begin{aligned} v_k(t,x,p) \;=\; \sup_{\alpha\in\mathcal{A}_{t,p}} \mathbb{E}\Big[\int_t^{t_1+mh} & f(X_s^{t,x,0})ds + c(X_{t_1+mh}^{t,x,0}, e_1) \\ & + v_{k-1}(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-)1_{\tau_{k+1}>t_1+mh} \\ & + v_k(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_- \cup (t_1+mh, \xi_{k+1}))1_{\tau_{k+1}=t_1+mh=t_k+h}\Big]. \end{aligned}$$

Now, from (6.43), if $t_1+mh = t_k+h$, we have $v_k(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-\cup(t_1+mh,\xi_{k+1})) \leq v_{k-1}(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-)$ for all ξ_{k+1} $\mathcal{F}_{t_1+mh}$-measurable valued in E. We then deduce

$$v_k(t,x,p) \;=\; \mathbb{E}\Big[\int_t^{t_1+mh} f(X_s^{t,x,0})ds + c(X_{t_1+mh}^{t,x,0}, e_1) + v_{k-1}(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-)\Big],$$

which is the required relation (6.39). (ii) The proof is analogous to (i), case 1, as if $\tau_i > t - mh$, then $\tau_i = +\infty$.

(iii) Fix $k = 1,\ldots,m$, $p = (t_i,e_i)_{1\leq i\leq k} \in \Theta_k \times E^k$, s.t. $t_k + h < t_1 + mh$, and $(t,x) \in \mathbb{T}_p^2(k)\times\mathbb{R}^d$. Then, for all $\alpha \in \mathcal{A}_{t,p}$, we have from (2.3), $X_s^{t,x,p,\alpha} = X_s^{t,x,0}$ for $t \leq s < t_1 + mh$, and $X_{t_1+mh}^{t,x,p,\alpha} = \Gamma(X_{t_1+mh}^{t,x,0}, e_1)$. Let $\alpha = (\tau_i,\xi_i)$ be some arbitrary element in $\mathcal{A}_{t,p}$, and set $\tau = \tau_{k+1}$, $\xi = \xi_{k+1}$. Observe that with $\theta = (t_1+mh)\wedge\tau$, we have a.s. either $k(\theta,\alpha) = k+1$, $p(\theta,\alpha) = p\cup(\tau,\xi)$ if $\tau < t_1+mh$ or $k(\theta,\alpha) = k-1$, $p(\theta,\alpha) = p_-$ if $\tau > t_1+mh$, or $k(\theta,\alpha) = k$, $p(\theta,\alpha) = p_-\cup(\tau,\xi)$ if $\tau = t_1+mh$. Hence, by applying (3.5) to some $\alpha = (\tau_i,\xi_i) \in \mathcal{A}_{t,p}$ s.t. $\tau_{k+1} > t_1+mh$ a.s. and with $\theta = t_1+mh$, we get the inequality (6.41). Furthermore, from (3.6), for all $\varepsilon > 0$, there exists $\alpha = (\tau_i,\xi_i) \in \mathcal{A}_{t,p}$ s.t. by setting $\tau = \tau_{k+1}$, $\xi = \xi_{k+1}$, and with $\theta = (t_1+mh)\wedge\tau$,

$$\begin{aligned} v_k(t,x,p) - \varepsilon \;\leq\; \mathbb{E}\Big[\int_t^{(t_1+mh)\wedge\tau} & f(X_s^{t,x,0})ds + v_{k+1}(\tau, X_\tau^{t,x,0}, p\cup(\tau,\xi))1_{\tau<t_1+mh} \\ & + c(X_{t_1+mh}^{t,x,0}, e_1)1_{t_1+mh\leq\tau} + v_{k-1}(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-)1_{t_1+mh<\tau} \\ & + v_k(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-\cup(\tau,\xi))1_{\tau=t_1+mh}\Big]. \end{aligned}$$

Now, we have $v_k(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-\cup(t_1+mh,\xi)) \leq v_{k-1}(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-)$ from (6.43). Since $(\tau,\xi) \in \mathcal{I}_t$, and ε is arbitrary, we deduce the required relation (6.42). $\square$

Proposition 6.4. *For all $k = 0, \dots, m$, v_k is continuous on $\mathcal{D}_k$. Moreover, for all $k = 1, \dots, m$, $p = (t_i, e_i)_{1 \le i \le k} \in \Theta_k \times E^k$, $x \in \mathbb{R}^d$,*

$$v_k((t_1 + mh)^-, x, p) \quad = \quad c(x, e_1) + v_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-).$$

Proof. We shall prove by forward induction on $n = m, \dots, N$ that **(Hk)(n)**, $k = 1, \dots, m(n)$, and **(H0)(n)** hold, where

(Hk)(n) v_k is continuous on $\mathcal{D}_k(n)$, and for all $p = (t_i, e_i)_{1 \le i \le k} \in \Theta_k(n) \times E^k$, $v_k((t_1 + mh)^-, x, p) = c(x, e_1) + v_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-)$, $x \in \mathbb{R}^d$.

(H0)(n) v_0 is continuous on $\mathcal{D}_0(n)$,

with the convention that **(Hk)(n)** is empty for $n = m$.

▶ **Initialization :** $n = m$. We know from proposition 4.1 that v_0 is a viscosity solution to (4.1) and (4.2) at step m. From lemma 6.4 we get $\overline{v_0}(T, x) = \underline{v_0}(T, x) = g(x)$. Together with the comparison principle at step $n = m$ in Proposition 6.3, we get $\overline{v_0} \le \underline{v_0}$ on $\mathcal{D}_0(m)$. This implies continuity of v_0 on $\mathcal{D}_0(m)$, i.e. **(H0)(m)** is satisfied.

▶ **Step** $n \to n+1$ **:** $n \in \{m, ..., N-1\}$**.** Suppose that **(Hk)(n)**, $k = 1, \dots, m(n)$, and **(H0)(n)** hold. Let us prove that **(Hk)(n+1)**, $k = 1, \dots, m(n+1)$, and **(H0)(n+1)** are satisfied.

• Take some $k = 1, \dots, m(n+1)$, and fix some arbitrary $x \in \mathbb{R}^d$ and $p = (t_i, e_i)_{1 \le i \le k} \in \Theta_k(n+1) \times \mathbb{E}^k$. Notice that $p_- \in \Theta_{k-1}(n) \times E^{k-1}$ so that $v_{k-1}(., ., p_-)$ is continuous on $\mathbb{T}_{p_-}(k-1) \times \mathbb{R}^d$ from step n. Here, to alleviate notations, we used the convention that $\mathbb{T}_{p_-}(k-1) = \mathbb{T}^n(0)$ if $k - 1 = 0$. We distinguish two cases :

⋆ *Case 1.* For some $\varepsilon > 0$, $\mathbb{T}^2_p(k) \cap [t_1 + mh - \varepsilon, t_1 + mh) = \emptyset$, i.e. $t_1 + mh \le t_k + h$ or $T - mh < t_1 + mh$ so that $[t_1 + mh - \varepsilon, t_1 + mh) \in \mathbb{T}^1_p(k)$. From (6.39) and (6.40), we then have for all $t \in [t_1 + mh - \varepsilon, t_1 + mh)$:

$$v_k(t, x, p) \quad = \quad \mathbb{E}\Big[\int_t^{t_1+mh} f(X_s^{t,x,0}) ds + c(X_{t_1+mh}^{t,x,0}, e_1) + v_{k-1}(t_1 + mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-) \Big].$$

By continuity of $v_{k-1}(t_1 + mh, ., p_-)$ (proved at step n), $\Gamma(., e_1)$, $c(., e_1)$, growth condition on f, c, Γ and v_{k-1}, we deduce with the dominated convergence theorem that $v_k((t_1 + mh)^-, x, p)$ exists and

$$v_k((t_1 + mh)^-, x, p) \quad = \quad c(x, e_1) + v_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-).$$

⋆ *Case 2.* $\mathbb{T}^2_p(k) = [t_k + h, t_1 + mh) \ne \emptyset$, i.e. $T - mh \ge t_1 + mh > t_k + h$ (this implies in particular that $k < (n + 1 - m) \wedge m$ and $m > 1$). From (6.41)-(6.42), we first prove that

$$\begin{aligned} & \overline{v_k}(t_1 + mh, x, p) \\ \le \;& \max\big[c(x, e_1) + v_{k-1}(t_1 + mh, x, p_-), \sup_{e \in E} \overline{v_{k+1}}(t_1 + mh, x, p \cup (t_1 + mh, e)) \big] \end{aligned} \tag{6.44}$$

Indeed, consider some sequence $(t_\varepsilon, x_\varepsilon, p_\varepsilon)_{\varepsilon > 0} \in \mathcal{D}_k$ converging to $(t_1 + mh, x, p)$ and such that $\lim_{\varepsilon \to 0} v_k(t_\varepsilon, x_\varepsilon, p_\varepsilon) = \overline{v_k}(t_1 + mh, x, p)$. For any $\varepsilon > 0$, one can find, by (6.42), some $(\hat{\tau}_\varepsilon, \hat{\xi}_\varepsilon) \in \mathcal{I}_{t_\varepsilon}$ s.t.

$$\begin{aligned} v_k(t_\varepsilon, x_\varepsilon, p_\varepsilon) \quad \le \quad & \mathbb{E}\Big[\int_{t_\varepsilon}^{(t_1^\varepsilon + mh) \wedge \hat{\tau}_\varepsilon} f(X_s^{t_\varepsilon, x_\varepsilon, 0}) ds + v_{k+1}(\hat{\tau}_\varepsilon, X_{\hat{\tau}_\varepsilon}^{t_\varepsilon, x_\varepsilon, 0}, p_\varepsilon \cup (\hat{\tau}_\varepsilon, \hat{\xi}_\varepsilon)) 1_{\hat{\tau}_\varepsilon < t_1^\varepsilon + mh} \\ & + \Big(c(X_{t_1^\varepsilon + mh}^{t_\varepsilon, x_\varepsilon, 0}, e_1^\varepsilon) + v_{k-1}(t_1^\varepsilon + mh, \Gamma(X_{t_1^\varepsilon + mh}^{t_\varepsilon, x_\varepsilon, 0}, e_1^\varepsilon), p_{\varepsilon -}) \Big) 1_{t_1^\varepsilon + mh \le \hat{\tau}_\varepsilon} \Big] \; + \varepsilon, \end{aligned}$$

where we denote $p_\varepsilon = (t_i^\varepsilon, e_i^\varepsilon)_{1\leq i\leq k}$ and $p_{\varepsilon-} = (t_i^\varepsilon, e_i^\varepsilon)_{2\leq i\leq k}$. By setting

$$G_\varepsilon \;=\; c(X^{t_\varepsilon,x_\varepsilon,0}_{t_1^\varepsilon+mh}, e_1^\varepsilon) + v_{k-1}(t_1^\varepsilon + mh, \Gamma(X^{t_\varepsilon,x_\varepsilon,0}_{t_1^\varepsilon+mh}, e_1^\varepsilon), p_{\varepsilon-}),$$

we rewrite the above inequality as

$$\begin{aligned} v_k(t_\varepsilon, x_\varepsilon, p_\varepsilon) \;\leq\; & \mathbb{E}\Big[\int_{t_\varepsilon}^{(t_1^\varepsilon+mh)\wedge\hat\tau_\varepsilon} f(X_s^{t_\varepsilon,x_\varepsilon,0})ds \;+\; G_\varepsilon \\ & + \Big(v_{k+1}(\hat\tau_\varepsilon, X^{t_\varepsilon,x_\varepsilon,0}_{\hat\tau_\varepsilon}, p_\varepsilon \cup (\hat\tau_\varepsilon, \hat\xi_\varepsilon)) - G_\varepsilon\Big)1_{\hat\tau_\varepsilon < t_1^\varepsilon+mh}\Big] \;+\varepsilon. \end{aligned} \tag{6.45}$$

Since $p_- \in \Theta_{k-1}(n)\times E^{k-1}$, we have $p_{\varepsilon-} \in \Theta_{k-1}(n)\times E^{k-1}$ for ε small enough. Hence, by continuity of v_{k-1} on $\mathcal{D}_{k-1}(n)$ (from part **1.**), continuity of Γ and c, and path-continuity of the flow $X_s^{t,x,0}$, we have

$$\lim_{\varepsilon\to 0} G_\varepsilon \;=\; G \;:=\; c(x,e_1) + v_{k-1}(t_1+mh, \Gamma(x,e_1), p_-) \qquad a.s. \tag{6.46}$$

Moreover, by compactness of E, the sequence $(\hat\xi_\varepsilon)_\varepsilon$ converges, up to a subsequence, to some ξ valued in E. We deduce that

$$\begin{aligned} & \limsup_{\varepsilon\to 0}\Big(v_{k+1}(\hat\tau_\varepsilon, X^{t_\varepsilon,x_\varepsilon,0}_{\hat\tau_\varepsilon}, p_\varepsilon\cup(\hat\tau_\varepsilon,\hat\xi_\varepsilon)) - G_\varepsilon\Big)1_{\hat\tau_\varepsilon<t_1^\varepsilon+mh} \\ \leq\; & \Big(\overline{v_{k+1}}(t_1+mh, x, p\cup(t_1+mh,\xi)) - G\Big)\limsup_{\varepsilon\to 0} 1_{\hat\tau_\varepsilon<t_1^\varepsilon+mh} \\ \leq\; & \Big(\sup_{e\in E}\overline{v_{k+1}}(t_1+mh, x, p\cup(t_1+mh,e)) - G\Big)\limsup_{\varepsilon\to 0} 1_{\hat\tau_\varepsilon<t_1^\varepsilon+mh} \quad a.s. \end{aligned} \tag{6.47}$$

From the linear growth condition on f, c, Γ, v_{k-1}, v_{k+1}, and estimate (2.8), we may use dominated convergence theorem and send ε to zero in (6.45) to obtain with (6.46)-(6.47) :

$$\begin{aligned} & \overline{v_k}(t_1+mh, x, p) \\ \leq\; & \mathbb{E}\Big[G + \Big(\sup_{e\in E}\overline{v_{k+1}}(t_1+mh,x,p\cup(t_1+mh,e)) - G\Big)\limsup_{\varepsilon\to 0}1_{\hat\tau_\varepsilon<t_1^\varepsilon+mh}\Big] \\ \leq\; & \max\Big[G, \sup_{e\in E}\overline{v_{k+1}}(t_1+mh,x,p\cup(t_1+mh,e))\Big], \end{aligned}$$

which is the required inequality (6.44).

We next show that

$$\sup_{e\in E}\overline{v_{k+1}}(t_1+mh,x,p\cup(t_1+mh,e)) \;\leq\; c(x,e_1) + v_{k-1}(t_1+mh,x,p_-). \tag{6.48}$$

Indeed, for any arbitrary $e\in E$, consider some sequence $(t_\varepsilon, x_\varepsilon, p_\varepsilon, e_\varepsilon)_{\varepsilon>0} \in \mathcal{D}_k\times E$ converging to (t_1+mh, x, p, e) and such that $\lim_{\varepsilon\to 0} v_{k+1}(t_\varepsilon, x_\varepsilon, p_\varepsilon\cup(t_\varepsilon,e_\varepsilon)) = \overline{v_{k+1}}(t_1+mh, x, p\cup(t_1+mh,e))$. For ε small enough, $t_\varepsilon + h \geq t_1^\varepsilon + mh$, and so from the DPP (6.39), we have :

$$\begin{aligned} v_{k+1}(t_\varepsilon, x_\varepsilon, p_\varepsilon\cup(t_\varepsilon,e_\varepsilon)) \;=\; & \mathbb{E}\Big[\int_{t_\varepsilon}^{t_1^\varepsilon+mh} f(X_s^{t_\varepsilon,x_\varepsilon,0})ds + c(X^{t_\varepsilon,x_\varepsilon,0}_{t_1^\varepsilon+mh}, e_1^\varepsilon) \\ & + v_k(t_1^\varepsilon+mh, \Gamma(X^{t_\varepsilon,x_\varepsilon,0}_{t_1^\varepsilon+mh}, e_1^\varepsilon), p_{\varepsilon-}\cup(t_\varepsilon,e_\varepsilon))\Big]. \end{aligned} \tag{6.49}$$

Since $p_- \in \Theta_{k-1}(n) \times E^{k-1}$, we have $p_{\varepsilon-} \in \Theta_{k-1}(n) \times E^{k-1}$ for ε small enough. Hence, by continuity of v_k on $\mathcal{D}_k(n)$, continuity and growth linear condition of f, Γ and c, and path-continuity of the flow $X_s^{t,x,0}$, we send ε to zero in (6.49) and get by the dominated convergence theorem

$$
\begin{aligned}
& \overline{v_{k+1}}(t_1 + mh, x, p \cup (t_1 + mh, e)) \\
= \;& c(x, e_1) + v_k(t_1 + mh, \Gamma(x, e_1), p_- \cup (t_1 + mh, e)).
\end{aligned} \tag{6.50}
$$

Moreover, from (6.43), we have $v_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-) \geq v_k(t_1 + mh, \Gamma(x, e_1), p_- \cup (t_1 + mh, e))$ for all $e \in E$. Plugging into (6.50), this proves (6.48).

Finally, we easily see from (6.41) that

$$
\underline{v_k}(t_1 + mh, x, p) \;\; \geq \;\; c(x, e_1) + v_{k-1}(t_1 + mh, x, p_-). \tag{6.51}
$$

Indeed, consider some sequence $(t_\varepsilon, x_\varepsilon, p_\varepsilon)_{\varepsilon > 0} \in \mathcal{D}_k$ converging to $(t_1 + mh, x, p)$ and such that $\lim_{\varepsilon \to 0} v_k(t_\varepsilon, x_\varepsilon, p_\varepsilon) = \underline{v_k}(t_1 + mh, x, p)$. From (6.41), we have in particular

$$
\begin{aligned}
v_k(t_\varepsilon, x_\varepsilon, p_\varepsilon) \;\; \geq \;\; & \mathbb{E}\Big[\int_{t_\varepsilon}^{t_1^\varepsilon + mh} f(X_s^{t_\varepsilon, x_\varepsilon, 0}) ds \\
& + c(X_{t_1^\varepsilon + mh}^{t_\varepsilon, x_\varepsilon, 0}, e_1^\varepsilon) + v_{k-1}(t_1^\varepsilon + mh, \Gamma(X_{t_1^\varepsilon + mh}^{t_\varepsilon, x_\varepsilon, 0}, e_1^\varepsilon), p_{\varepsilon-}) \Big].
\end{aligned}
$$

By continuity and linear growth condition of v_{k-1}, Γ, c, f, and estimate (2.8), we get (6.51) by the dominated convergence theorem, and sending ε to zero in the above inequality.

Hence, the inequalities (6.44)-(6.48)-(6.51) prove that $v_k((t_1 + mh)^-, x, p)$ exists and is equal to :

$$
\begin{aligned}
v_k((t_1 + mh)^-, x, p) \;\; &= \;\; \overline{v_k}(t_1 + mh, x, p) \;\; = \;\; \underline{v_k}(t_1 + mh, x, p) \\
&= \;\; c(x, e_1) + v_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-).
\end{aligned} \tag{6.52}
$$

We have then proved that (6.52) holds for all $k = 1, \ldots, m$, $p = (t_i, e_i)_{1 \leq i \leq k} \in \Theta_k(n+1) \times E^k$, and $x \in \mathbb{R}^d$.

• We know from Proposition 4.1 that the family of value functions v_k, $k = 0, \ldots, m(n+1)$, is a viscosity solution to (4.1)-(4.2), in particular at step $n+1$. We also recall from Lemma 6.4 that $\overline{v_0}(T, x) = \underline{v_0}(T, x) = g(x)$. Together with (6.52), and the comparison principle at step $n+1$ in Proposition 6.3, this proves $\overline{v_k} \leq \underline{v_k}$ on $\mathcal{D}_k(n+1)$. This implies the continuity of v_k on $\mathcal{D}_k(n+1)$, $k = 0, \ldots, m(n+1)$, and so **(Hk)(n+1)**, $k = 1, \ldots, m(n+1)$, and **(H0)(n+1)** are stated.

► The proof is completed at step N by recalling that $\Theta_k(N) = \Theta_k$, $\mathcal{D}_k(N) = \mathcal{D}_k$, for $k = 0, \ldots, m(N) = m$. □

6.5 Proof of Theorem 4.1

In view of the results proved in paragraphs 6.2 and 6.4, it remains to prove the uniqueness result of Theorem 4.1. Let us then consider another family w_k, $k = 0, \ldots, m$ of viscosity solutions to (4.1)-(4.2), satisfying growth condition (2.9), and boundary data (4.4)-(4.5) : for $k = 1, \ldots, m$, $p = (t_i, e_i)_{1 \leq i \leq k} \in \Theta_k \times E^k$, $x \in \mathbb{R}^d$,

$$
w_k((t_1 + mh)^-, x, p) \;\; = \;\; c(x, e_1) + w_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-). \tag{6.53}
$$

and

$$w_0(T^-, x) \quad = \quad g(x), \quad x \in \mathbb{R}^d. \tag{6.54}$$

We shall prove by forward induction on $n = m, \ldots, N$ that $v_k = w_k$ on $\mathcal{D}_k(n)$.

▶ **Initialization :** $n = m$. Relations (4.5), (6.54) and Proposition 6.3 at step $n = m$ show that $v_0 = w_0$ on $\mathcal{D}_0(m)$.

▶ **Step** $n \to n+1$. Suppose that $v_k = w_k$ on $\mathcal{D}_k(n)$, $k = 0, \ldots, m(n)$. For any $k \geq 1$, $p = (t_i, e_i)_{1 \leq i \leq k} \in \Theta_k(n+1) \times E^k$, we notice that $p_- \in \Theta_{k-1}(n) \times E^{k-1}$. Hence $v_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-) = w_{k-1}(t_1 + mh, \Gamma(x, e_1), p_-)$, $x \in \mathbb{R}^d$, and so from (4.4), (6.53), we have

$$v_k((t_1 + mh)^-, x, p) \quad = \quad w_k((t_1 + mh)^-, x, p).$$

We already know that $v_0(T^-, x) = w_0(T^-, x)$ $(= g(x))$. Therefore, from the comparison principle at step $n+1$ in Proposition 6.3, we deduce that $u_k = w_k$ on $\mathcal{D}_k(n+1)$, $k = 0, \ldots, m(n)$. Finally, the proof is completed since $\mathcal{D}_k(N) = \mathcal{D}_k$.

Chapter 5

A numerical algorithm for impulse control problems with execution delay

In this chapter we describe a numerical algorithm to solve impulse control problems with execution delay on finite horizon. In this problem, the family of value functions is characterized by a family of variational inequalities. The main contribution of our work is a general algorithm which enables to calculate the solutions of this sequence of variational inequalities in a correct order. Then, we approximate the solution of each equation with a finite differences scheme. We prove the convergence of this scheme by the method of [10], in the framework of viscosity solutions. Finally, we give a concrete financial illustration with several numerical results.

Key words : Impulse control, finite differences, viscosity solutions, execution delay.

1 Introduction

In this chapter, we describe the numerical procedure for computing solutions of the delay control problem of the former chapter. It fits in the general framework of impulse control problems. For an overview of these kind of problems, the reader may refer to [13] and [56]. The numerical computations of variational inequalities arising in impulse control and optimal stopping problems have been studied by many authors, for instance in [21],[22] and [8]. Other authors, like in [44] and [67] studied equations arising from singular control. These methods can be roughly divided in two kinds. The probabilistic ones involve most of the time Monte Carlo simulations of backward stochastic differential equations (see [19] for instance). The analytic ones, which we will use, involve the approximate resolution of a discretized PDE on the whole domain. More precisely, we will consider finite difference methods. There are many theoretical frameworks to derive properties such as the convergence rate of a finite difference scheme. The choice of a correct framework mainly depends of the regularity of the function we try to approximate. In our case, we can not formulate any regularity property for the value function of our problem, excepted its continuity. Therefore, we have to work with the viscosity solutions theory. The reader can refer to [27] for a general introduction to this concept. The convergence of numerical scheme for viscosity solution has been proved in a very general way in [10]. The rate of convergence has been subject to a wide number of studies, see [9] for instance. However, here, we will only consider the convergence property, leaving its rate to further research. In the former chapter, keeping the same notations, we proved that the family of value functions of the delay control problem is the unique viscosity solution of the equation:

$$F^1(t,x,Dv_k,D^2v_k) = 0 \quad \text{on} \quad \mathcal{D}_k^1, \quad k = 1,\dots,m, \tag{1.1}$$

and:

$$F^2(t,x,Dv_k,D^2v_k,v_{k+1}) = 0 \quad \text{on} \quad \mathcal{D}_k^2, \ \ k = 0,\dots,m-1, \tag{1.2}$$

with:

$$\begin{aligned} F^1(t,x,Dv_k,D^2v_k) &= -\frac{\partial v_k}{\partial t}(t,x,p) - \mathcal{L}v_k(t,x,p) - f(x) \\ F^2(t,x,Dv_k,D^2v_k,v_{k+1}) &= \min\Big\{ -\frac{\partial v_k}{\partial t}(t,x,p) - \mathcal{L}v_k(t,x,p) - f(x)\ , \\ & \qquad v_k(t,x,p) - \sup_{e\in E} v_{k+1}(t,x,p\cup(t,e))\Big\} \end{aligned}$$

The value functions satisfy linear growth and the following boundary conditions:
(i) For $k = 1,\dots,m$, $p = (t_i,e_i)_{1\le i\le k} \in \Theta_k \times E^k$, $x \in \mathbb{R}^d$, $v_k((t_1+mh)^-,x,p)$ exists and :

$$v_k((t_1+mh)^-,x,p) \;=\; c(x,e_1) + v_{k-1}(t_1+mh,\Gamma(x,e_1),p_-). \tag{1.3}$$

(ii) For all $x \in \mathbb{R}^d$, $v_0(T^-,x)$ exists and is equal to :

$$v_0(T^-,x) \;=\; g(x). \tag{1.4}$$

The difficulty with this formulation is that there is a mixed dependence between v_k and v_{k+1}. Indeed, to compute the solution v_k of PDE (1.2), we need to know the values of v_{k+1}.

On the other hand, to compute v_{k+1}, we need the values of v_k to obtain the boundary condition (1.3). This work gives a method to handle these dependencies. To this end, we will use an iterative algorithm described in section 2 that reduces this problem to solving a sequence of variational inequalities and linear PDEs. Then, we will use classical algorithms for obstacle problems to solve each of these equations. Indeed, supposing that v_{k+1} is known, equation (1.2) can be interpreted as an optimal stopping problem. Therefore, one can use some finite difference method developed for optimal stopping problems, such as the Howard algorithm for implicit schemes (see [49] for instance). However, a major difference with respect to the standard case is that the obstacle v_{k+1} and the terminal conditions are endogenous, and is approximated jointly with v_k. This is why we can not use standard theorems to prove convergence directly.
In the rest of this work, we will use the well known condition for finite difference schemes :

Assumption 1.1. *(i) $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz continuous functions. (ii) The matrix $\sigma(x)\sigma^T(x)$ is strictly diagonally dominant for all $x \in \mathbb{R}^d$.*

This work is organized as follows. In section 2, we present the general algorithm to compute the value functions by solving variational inequalities. In section 3, we present a discrete scheme one can use to solve variational inequalities of the form (1.1), (1.2). In section 4, we give some properties of this scheme, and we prove its convergence. In section 5, we introduce an impulse control problem without delay, in order to compare its value function with the delay problem case. Finally, in section 6, we consider an example of financial application of delayed control problem for optimal investment, and for the pricing of a call option with an illiquid underlying. We describe the problem and give some some numerical results.

2 General algorithm to compute the value function

We first make the following observation: Let us denote by H_0 the function defined on $[0, T-mh] \times \mathbb{R}^d$ by

$$H_0(t,x) \quad = \quad \sup_{e\in E} v_1(t,x,(t,e)).$$

And for $t \in (T-mh, T]$, we denote:

$$H_0(t,x) \quad = \quad \mathbb{E}\Big[\int_t^T f(X_s^{t,x,0})ds + g(X_T^{t,x,0})\Big], \quad (t,x) \in (T-mh, T].$$

This function H_0 clearly satisfies the linear PDE : $-\dfrac{\partial H_0}{\partial t} - \mathcal{L}H_0 - f = 0$ on $(T-mh, T)$ together with the terminal condition $H_0(T^-, x) = g(x) = v_0(T^-, x)$. Hence, with (1.1) and (1.4) for $k = 0$, this shows that

$$v_0(t,x) \quad = \quad H_0(t,x), \qquad (t,x) \in (T-mh, T] \times \mathbb{R}^d, \tag{2.1}$$

Moreover, from the PDE (1.2) for $k = 0$, and a standard uniqueness result for the corresponding free-boundary problem, we may also represent v_0 as the solution to the optimal stopping problem :

$$v_0(t,x) \quad = \quad \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}[H_0(\tau, X_\tau^{t,x,0})], \quad (t,x) \in [0,T] \times \mathbb{R}^d, \tag{2.2}$$

where $\mathcal{T}_{t,T}$ denotes the set of stopping times τ valued in $[t,T]$. Hence, the value function v_0 is completely determined once we can compute v_1.

We show how one can compute $v_k(.,.,p)$ on $\mathbb{T}_p(k)\times\mathbb{R}^d$ for all $p\in\Theta_k\times E^k$, $k=1,\ldots,m$ and v_0 on $[0,T]\times\mathbb{R}^d$.

For $k=1,\ldots,m$, and any $n\geq 1$, we denote :

$$
\begin{aligned}
\Theta_k(n) &= \Big\{t^{(k)}=(t_i)_{1\leq i\leq k}\in\Theta_k\ : t_1>T-nh\Big\},\\
N &= \inf\{n\geq 1\ : T-nh<0\},
\end{aligned}
$$

so that $\Theta_k(n)$ is strictly included in $\Theta_k(n+1)$ for $n=1,\ldots,N-1$, and $\Theta_k(N)=\Theta_k$. We also denote for $k=0$, and $n\geq 1$, $\mathbb{T}^n(0)=(T-nh,T]\cap[0,T]$ so that $\mathbb{T}^n(0)=(T-nh,T]$ is increasing with $n=1,\ldots,N-1$, and $\mathbb{T}^N(0)=[0,T]$. We assumed $T-mh\geq 0$ to avoid trivialities so that $N>m$. We denote for $k=0,\ldots,(n-m)\wedge m$, and $n=m,\ldots,N$,

$$
\begin{aligned}
\mathcal{D}_k(n) &= \Big\{(t,x,p)\in\mathcal{D}_k\ : p\in\Theta_k(n)\times E^k\Big\},\\
\mathcal{D}_k^i(n) &= \mathcal{D}_k(n)\cap\mathcal{D}_k^i\ =\ \Big\{(t,x,p)\in\mathcal{D}_k(n)\ : t\in\mathbb{T}_p^i(k)\Big\},\ \ i=1,2,
\end{aligned}
$$

with the convention that $\mathcal{D}_0(n)=\mathbb{T}^n(0)\times\mathbb{R}^d$, so that $\mathcal{D}_k(n)$ is strictly included in $\mathcal{D}_k(n+1)$ for $n=1,\ldots,N-1$, and $\mathcal{D}_k(N)=\mathcal{D}_k$. We shall compute v_k on $\mathcal{D}_k(n)$, $k=0,\ldots,m$, by forward induction on $n=m,\ldots,N$ and backward induction on k.

▶ **Initialization phase : $n=m$.** From (1.4) and (2.1), we know the values of v_0 on $\mathcal{D}_0(m)$:

$$
v_0(t,x) = \mathbb{E}\Big[\int_t^T f(X_s^{t,x,0})ds+g(X_T^{t,x,0})\Big].
$$

▶ **Step $n\to n+1$ for $n\in\{m,\ldots,N-1\}$.** We denote $m(n)=(n-m)\wedge m$ the maximum number of pending orders at step n. Suppose we know the values of v_k on $\mathcal{D}_k(n)$, $k=0,\ldots,m(n)$. In order to determine v_k on $\mathcal{D}_k(n+1)$, $k=0,\ldots,m(n+1)$, it suffices to compute $v_k(.,.,p)$ on $\mathbb{T}_p(k)\times\mathbb{R}^d$ for all $p\in\Theta_k(n+1)\times E^k$, $k=1,\ldots,m(n+1)$, and v_0 on $\mathbb{T}^{n+1}(0)\times\mathbb{R}^d$. We shall argue by backward induction on $k=m(n+1),\ldots,0$.

- Let $k=m(n+1)$, and take some arbitrary $p=(t_i,e_i)_{1\leq i\leq m(n+1)}\in\Theta_{m(n+1)}(n+1)\times E^{m(n+1)}$. Recall that $\mathbb{T}_p^2(m(n+1))$ is empty so that $\mathbb{T}_p(m(n+1))=\mathbb{T}_p^1(m(n+1))$ $=[t_{m(n+1)},t_1+mh)$. From (1.3) for $k=m$, we have $v_{m(n+1)}((t_1+mh)^-,x,p)=c(x,e_1)+v_{m(n+1)-1}(t_1+mh,\Gamma(x,e_1),p_-)$ for all $x\in\mathbb{R}^d$, which is known from step n since either $p_-\in\Theta_{m(n+1)-1}(n)\times E^{m(n+1)-1}$ when $m(n+1)>1$, or $t_1+mh\in\mathbb{T}^n(0)$ when $m(n+1)-1=0$. We then solve $v_{m(n+1)}(.,.,p)$ on $\mathbb{T}_p^1(m(n+1))\times\mathbb{R}^d$ from (1.1) for $k=m(n+1)$, which gives :

$$
\begin{aligned}
v_{m(n+1)}(t,x,p) &= \mathbb{E}\Big[\int_t^{t_1+mh} f(X_s^{t,x,0})ds+c(X_{t_1+mh}^{t,x,0},e_1)\\
&\qquad\qquad + v_{m(n+1)-1}(t_1+mh,\Gamma(X_{t_1+mh}^{t,x,0},e_1),p_-)\Big].
\end{aligned}
$$

  We have then computed the value of $v_{m(n+1)}(.,.,p)$ on $\mathbb{T}_p(m(n+1))\times\mathbb{R}^d$.

- From $k+1 \to k$ for $k = m(n+1)-1, \ldots, 1$. (This step is empty when $m(n+1) = 1$). Suppose we know the values of $v_{k+1}(.,.,p)$ on $\mathbb{T}_p(k+1) \times \mathbb{R}^d$ for all $p \in \Theta_{k+1}(n+1) \times E^{k+1}$. Take now some arbitrary $p = (t_i, e_i)_{1 \le i \le k} \in \Theta_k(n+1) \times E^k$. We shall compute $v_k(.,.,p)$ successively on $\mathbb{T}_p^2(k) \times \mathbb{R}^d$ (if it is not empty) and then on $\mathbb{T}_p^1(k) \times \mathbb{R}^d$, and we distinguish the two cases :

  (i) $\mathbb{T}_p^2(k) = \emptyset$. This means $t_k + h \ge t_1 + mh$ or $t_k + h > T - mh$, and so $\mathbb{T}_p(k) = \mathbb{T}_p^1(k) = [t_k, t_1 + mh)$. We then compute $v_k(.,.,p)$ on $\mathbb{T}_p(k) \times \mathbb{R}^d$ as above for $k = m$:

$$\begin{aligned} v_k(t,x,p) &= \mathbb{E}\Big[\int_t^{t_1+mh} f(X_s^{t,x,0}) ds + c(X_{t_1+mh}^{t,x,0}, e_1) \\ &\qquad + v_{k-1}(t_1+mh, \Gamma(X_{t_1+mh}^{t,x,0}, e_1), p_-)\Big], \end{aligned}$$

  where the r.h.s. is known from step n since either $p_- \in \Theta_{k-1}(n) \times E^{k-1}$ when $k > 1$, or $t_1 + mh \in \mathbb{T}^n(0)$ when $k - 1 = 0$.

  (ii) $\mathbb{T}_p^2(k) \neq \emptyset$. This means $t_k + h < t_1 + mh$ and $t_k + h \le T - mh$, so $\mathbb{T}_p^1(k) = [t_k, t_k + h) \cup \Big([t_k, t_k+h) \cap (T - mh, T)\Big)$, $\mathbb{T}_p^2(k) = [t_k + h, t_1 + mh) \cap [0, T - mh]$. For all $(t,x) \in \mathbb{T}_p^2(k) \times \mathbb{R}^d$, and $e \in E$, we have $p' = p \cup (t,e) \in \Theta_{k+1}(n+1) \times E^{k+1}$, and $(t,x) \in \mathbb{T}_{p'}(k+1) \times \mathbb{R}^d$. Hence, from the induction hypothesis at order $k+1$, we know the value of the function :

$$H_{k,p}(t,x) = \sup_{e \in E} v_{k+1}(t,x,p \cup (t,e)), \quad (t,x) \in \mathbb{T}_p^2(k) \times \mathbb{R}^d.$$

  We also know from step n the value of the function :

$$G_{k,p}(x) = c(x,e_1) + v_{k-1}(t_1 + mh, \Gamma(x,e_1), p_-), \quad x \in \mathbb{R}^d.$$

  Then, from the PDE (1.2) and the terminal condition (1.3) at k, we compute $v_k(.,.,p)$ on $\mathbb{T}_p^2(k) \times \mathbb{R}^d$ as the solution to an optimal stopping problem with obstacle $H_{k,p}$ and terminal condition $G_{k,p}$:

$$\begin{aligned} v_k(t,x,p) &= \sup_{\tau \in \mathcal{T}_{t,t_1+mh}} \mathbb{E}[H_{k,p}(\tau, X_\tau^{t,x,0}) 1_{\tau < t_1+mh} \\ &\qquad + G_{k,p}(X_{t_1+mh}^{t,x,0}) 1_{\tau = t_1+mh}], \quad (t,x) \in \mathbb{T}_p^2(k) \times \mathbb{R}^d. \end{aligned}$$

  In particular, by continuity of $v_k(.,.,p)$ on $\mathbb{T}_p(k)$, we know the value of $\lim_{t \nearrow t_k+h} v_k(t,x,p) = v_k(t_k + h, p)$. We then compute $v_k(.,.,p)$ on $\mathbb{T}_p^1(k) \times \mathbb{R}^d$ from (1.1) :

$$v_k(t,x,p) = \mathbb{E}\Big[\int_t^{t_k+h} f(X_s^{t,x,0}) ds + v_k(t_k+h, X_{t_k+h}^{t,x,0}, p)\Big].$$

  We have then computed the value of $v_k(.,.,p)$ on $\mathbb{T}_p(k) \times \mathbb{R}^d$.

- From $k = 1 \to k = 0$. From the above item, we know the value of $v_1(.,.,p)$ on $\mathbb{T}_p(1) \times \mathbb{R}^d$ for all $p \in \Theta_1(n+1) \times E$. Hence, we know the value of :

$$H_0(t,x) = \sup_{e \in E} v_1(t,x,(t,e)), \quad \forall\, (t,x) \in \mathbb{T}^{n+1}(0) \times \mathbb{R}^d.$$

From (2.2), we then compute v_0 on $\mathbb{T}^{n+1}(0) \times \mathbb{R}^d$ as an optimal stopping problem with obstacle F_1.

We have then calculated $v_k(.,.,p)$ on $\mathbb{T}_p(k) \times \mathbb{R}^d$ for all $p \in \Theta_k(n+1) \times E^k$ and v_0 on $\mathbb{T}^{n+1}(0) \times \mathbb{R}^d$, and step $n+1$ is stated. Finally, at step $n = N$, the computation of the value functions is completed since $\mathcal{D}_k(N) = \mathcal{D}_k$, $k = 0, \ldots, m$.

3 The discrete scheme

In this section, we describe the numerical scheme we use for equations (1.1) and (1.2). We suppose that the scheme can be calculated on the whole space $\mathcal{D}_k$, $k = 0, \ldots, m$. We suppose that the derivatives of the function on $\mathbb{R}^d$ are calculated with a space step $\delta_x = (\delta_1, ..., \delta_n)$, δ_i being the space step for direction f_i, where $(f_i)_{i=1\ldots d}$ is the canonical base of $\mathbb{R}^d$. We use a uniform time step δ_t. We also need to discretize with respect to p. This variable is of the form $p = (t_i, e_i)_{i \in \{1,\ldots,k\}}$ with each $e_i \in E$. Here we will consider the case $E = [e_{min}, e_{max}]$. We choose a step $\delta_{t_p} = \delta_t$ for the discretization of the times at which the orders are passed, and a step δ_e for the values of these orders. Note that one could use a step $\delta_{t_p} = n\delta_t$ for some positive integer n, using linear interpolation to obtain the missing values, and the scheme would work as well. Nevertheless, we choose $n = 1$ for the sake of simplicity. To obtain a concise notation, we will write:

$$\delta = (\delta_t, \delta_x, \delta_{t_p}, \delta_e)$$

Now, we need to reduce the problem to a bounded space. This will be done with assumptions 3.2 and 3.3 below. We define the set:

$$A = \{x = (x_1, \ldots, x_d) \; : \; x_i \in (x_{i,min}, x_{i,max}) \forall i\} \, .$$

Intuitively, these assumptions will constrain the process X_t to be confined in the space A. With these assumptions, while still considering the equation on $\mathbb{R}^d$, one can localize the discrete scheme to a bounded space:

$$A^{\delta_x} = \{x = (x_1, \ldots, x_d) \; : \; x_i \in (x_{i,min} - \delta_{x_i}, x_{i,max} + \delta_{x_i}) \forall i\} \, .$$

Therefore, considering a given point $x \in A^{\delta_x}$, the scheme will involve the following discrete grid:

$$\Omega_{x,\delta_x} = \left\{ \tilde{x} = (x_1 + n_1\delta_1, \ldots, x_d + n_d\delta_d) \in A^{\delta_x}, \; (n_1, \ldots, n_d) \in \mathbb{Z}^d \right\} .$$

At last, we will discretize the space E, to obtain a computable scheme. To this end, we introduce the set:

$$E^{\delta_e} = \{e = e_{min} + n\delta_e \; : \; e \in E, \; n \in \mathbb{N}\}$$

3.1 Boundary conditions

In this section, we state the boundary conditions we will use in this framework. We consider a function Ψ_k defined on $\mathcal{D}_k$. First of all, there are the terminal conditions that were described in the preceding section. The first one is:

$$\Psi_0^\delta(t, x) = g(x) \text{ for all } (t, x) \in [T - \delta_t, T) \times \mathbb{R}^d \tag{3.1}$$

Secondly, we must take into account condition (1.3) which gives, for any $k = 1, \dots, m$, and recalling the notations of the former chapter :

$$\Psi_k^\delta(t,x,p) = \Psi_{k-1}^\delta(t_1 + mh, \Gamma(x,e_1), p^-) + c(x,e_1), \tag{3.2}$$

for all $(x,p) \in \mathbb{R}^d \times \Theta_k$, and $t \in [t_1 + mh - \delta_t, t_1 + mh)$.
In practice, if $\Psi_{k-1}^\delta(t_1+mh, \Gamma(x,e_1), p^-)$ is not computed, on can use a linear interpolation. Nevertheless, we will consider that the scheme is calculated on $\mathcal{D}_k$, $k = 0, \dots, m$. Therefore we will not have this problem. Finally, we suppose that we do not need conditions on the boundary of Ω_{x,δ_x}. To have good framework, we suppose that:

Assumption 3.2. *For all $x \in \mathbb{R}^d$, if $x_i \leq x_{i,min}$ or $x_i \geq x_{i,max}$ for some $i \in \{1 \dots d\}$, then $\big(\mu(x)\big)_i = 0$ and $\big(\mu(x)\big)_{i,j} = 0$ for all $j \in \{1 \dots d\}$.*

With this assumption, we do not need to specify any conditions at the boundaries of A^{δ_x}. Finally we need an assumption to be sure that we can reduce the problem to the domain A^{δ_x}.

Assumption 3.3. *Any set $B \subset \mathbb{R}^d$ such that:*

$$\{x = (x_1, \dots, x_d) : x_i \in [x_{i,min}, x_{i,max}] \ \forall i \ \} \subset B$$

is stable with respect to Γ, that is for all $e \in E$:

$$\Gamma(B,e) \subset B$$

3.2 Discretization of the operators

We recall the classical space discretization of linear PDEs with finite difference schemes:

$$L^{\delta_x}(t,x,p,\Psi_k) = f(x) + \sum_{i=1}^d \mu_i(x) \partial_x^{i,sign(\mu_i)} \Psi_k + \sum_{(i,j) \in \{1..d\}^2} \frac{1}{2} (\sigma\sigma^T(x))_{i,j} \partial_{xx}^{i,j} \Psi_k$$

with the first order differential operators:

$$\partial_x^{i,+}\Psi_k = \frac{\Psi_k(t, x + \delta_i f_i, p) - \Psi_k(t,x,p)}{\delta_i}$$
$$\partial_x^{i,-}\Psi_k = \frac{\Psi_k(t,x,p) - \Psi_k(t, x - \delta_i f_i, p)}{\delta_i},$$

and the second order differences:

$$\partial_{xx}^{i,i,+} = \partial_{xx}^{i,i,-} = \frac{1}{\delta_i^2}\Big(2\Psi_k(t,x,p) - \Psi_k(t, x + \delta_i f_i, p) - \Psi_k(t, x - \delta_i f_i, p)\Big)$$

$$\partial_{xx}^{i,j,+} = \frac{1}{2\delta_i\delta_j}\Big(2\Psi_k(t,x,p) + \Psi_k(t, x + \delta_i f_i + \delta_j f_j, p) + \Psi_k(t, x - \delta_i f_i - \delta_j f_j, p)$$
$$-\Psi_k(t, x + \delta_j f_j, p) - \Psi_k(t, x - \delta_j f_j, p) - \Psi_k(t, x + \delta_i f_i, p) - \Psi_k(t, x - \delta_i f_i, p)\Big)$$

$$\partial_{xx}^{i,j,-} = \frac{1}{2\delta_i\delta_j}\Big(2\Psi_k(t,x,p) + \Psi_k(t, x + \delta_i f_i - h_j f_j, p) + \Psi_k(t, x - \delta_i f_i + \delta_j f_j, p)$$
$$-\Psi_k(t, x + \delta_j f_j, p) - \Psi_k(t, x - \delta_j f_j, p) - \Psi_k(t, x + \delta_i f_i, p) - \Psi_k(t, x - \delta_i f_i, p)\Big).$$

3.3 Time discretization: the linear case

For the linear equation (1.1), one can perform discretization in time of F^1 using classical θ-scheme as follows:

$$\begin{aligned}
\frac{S^{1,\delta}((t,x,p),\Psi_k(t,x,p),\Psi_k)}{\delta_t} &= \\
\frac{\Psi_k(t,x,p)-\Psi_k(t+\delta_t,x,p)}{\delta_t} &- \theta L^\delta(t,x,p,\Psi_k)-(1-\theta)L^\delta(t+\delta_t,x,p,\Psi_k)=0 \\
\Leftrightarrow \Psi_k(t,x,p)-\delta_t\theta L^\delta(t,x,p,\Psi_k) &- \Psi_k(t+\delta_t,x,p)-\delta_t(1-\theta)L^\delta(t+\delta_t,x,p,\Psi_k)=0.
\end{aligned}$$

This leads to a sequence of linear systems on Ω_{x,δ_x}. We denote $N = card(\Omega_{x,\delta_x})$. We suppose we order the elements of Ω_{x,δ_x}, so that:

$$\Omega_{x,\delta_x} = \left\{x^i, i\in\{1\dots N\}\right\}.$$

In this framework, for fixed p and t, we write:

$$X = \begin{pmatrix} \Psi_k(t,x^1,p) \\ \vdots \\ \Psi_k(t,x^N,p) \end{pmatrix}$$

Then the scheme can be written as:

$$A^0X - b^0 = 0, \tag{3.3}$$

where $X\in\mathbb{R}^N$ contains the values of $\Psi_k(t,x,p)$ for all $x\in\Omega_{x,\delta_x}$. The vector $b^0\in\mathbb{R}^N$, and the $N\times N$ squared matrix A^0 are functions of $\Psi(t+\delta_t,.,p)$, as described above. We have the following property of A, which is important for stability and monotonicity issues.

Proposition 3.5. *If $\delta_i=\delta_j$ for all $(i,j)\in\{1...d\}$, and if $(\sigma\sigma^T(x))$ is strictly diagonally dominant for all x, then the matrix A^0 is strictly diagonally dominant.*

This leads to the monotonicity and stability properties of the scheme for the case of the implicit scheme $\theta=1$ that we will consider in the rest of this work.

3.4 The non linear case

Now that we exposed the discrete scheme for the linear equation (1.1), we consider equation (1.2). We assume that Ψ_{k+1} has already been calculated. That equation can be discretized as follows:

$$\begin{aligned}
\frac{S_k^{2,\delta}((t,x,p),\Psi_k(t,x,p),\Psi_k,\Psi_{k+1})}{\delta_t} &= \\
\min\Bigg\{\frac{\Psi_k(t,x,p)-\Psi_k(t+\delta_t,x,p)}{\delta_t} &- \theta L^\delta(t,x,p,\Psi_k)-(1-\theta)L^\delta(t+\delta_t,x,p,\Psi_k), \\
\Psi_k(t,x,p) &- \sup_{e\in E^{\delta_e}}\left\{\Psi_{k+1}(t,x,p\cup(t,e))\right\}\Bigg\}=0.
\end{aligned}$$

It can be written in the following form:

$$\min_{\alpha\in\{0,1\}^N}\{A^\alpha X-b^\alpha\}=0. \tag{3.4}$$

the parameter α controls each line of A and b. It acts as follows:

- When $\alpha_i = 0$, the i-th line and A^α and b^α are equal to the i-th lines of A^0 and b^0 in the linear case (3.3).
- When $\alpha_i = 1$, the i-th line and A^α is equal to the i-th line of the identity matrix, and $b_i^\alpha = \sup_{e \in E^{\delta_e}} \{\Psi_{k+1}(t, x_i, p \cup (t, e))\}$.

3.5 Howard algorithm

To solve equation (3.4), we use the Howard algorithm, developed in [46]. It is an iterative algorithm on the controls α. It can be described as follows:

- Step 1: For some arbitrary chosen α_0, calculate the solution X^1 of the system:
$$A^{\alpha_0} X^1 - b^{\alpha_0} = 0$$
- Step $2n$: Calculate the new control α_n by:
$$\alpha_n = \arg \min_{\alpha \in \{0,1\}^N} (A^{\alpha_n} X^n - b^{\alpha_n})$$
- Step $2n+1$: Calculate the new value X^{n+1} of X by solving the linear system:
$$A^{\alpha_{n+1}} X^{n+1} - b^{\alpha_{n+1}} = 0$$
and stop the algorithm if $X^{n+1} = X^n$. Else proceed to step $2(n+1)$.

If the matrix A^α is diagonally dominant for all α this algorithm stops after a finite number of iterations n, see [49] for details. It means that in this case we obtain:
$$A^{\alpha_{n+1}} X^{n+1} - b^{\alpha_{n+1}} = A^{\alpha_{n+1}} X^n - b^{\alpha_{n+1}}$$
therefore, with the definition of α_{n+1} we obtain:
$$\min_{\alpha \in \{0,1\}^N} (A^\alpha X^n - b^\alpha) = 0.$$
Which is the solution of our problem.

4 Convergence of the discrete scheme

In this section, we will prove the convergence of the solution of the implicit scheme $\theta = 1$ to the value function. As we deal with viscosity solutions, we will use the method of [10]. This method can be applied with PDE satisfying a strong comparison principle for viscosity solutions. It states that a stable, monotone and consistent scheme necessarily converges to the correct solution. In our case, an additional difficulty comes from the terminal conditions which are not stated in the viscosity sense. They have to be stated as a restriction on the class of functions on which the comparison principle holds. Therefore, to prove convergence of the algorithm, we must first prove that the limit of the solutions of the scheme satisfies the terminal conditions. For the sake of conciseness, we will denote, by a misuse of notations:
$$\begin{aligned} F((t,x,p), Dv_k, D^2 v_k, v_{k+1}) &= 1_{(t,x,p) \in \mathcal{D}_k^1} F^1((t,x,p), Dv_k, D^2 v_k) \\ &\quad + 1_{(t,x,p) \in \mathcal{D}_k^2} F^2((t,x,p), Dv_k, D^2 v_k, v_{k+1}), \end{aligned}$$

even when v_{k+1} is not defined. Remark that in that case, $(t,x,p) \in \mathcal{D}_k^1$, hence there is no ambiguity. We will also denote, the same way:

$$\begin{aligned} S^\delta((t,x,p),\Psi_k,\Psi_k(t,x,p),\Psi_{k+1}) &= 1_{(t,x,p)\in\mathcal{D}_k^1} S^{1,\delta}((t,x,p),\Psi_k,\Psi_k(t,x,p)) \\ &\quad + 1_{(t,x,p)\in\mathcal{D}_k^2} S^{2,\delta}((t,x,p),\Psi_k,\Psi_k(t,x,p),\Psi_{k+1}). \end{aligned}$$

Finally, we say that a family of functions Ψ_k, $k = 0,\ldots,m$ is a supersolution (resp subsolution) of the discrete equation:

$$S^\delta((t,x,p),\Psi_k,\Psi_k(t,x,p),\Psi_{k+1}) = 0 \tag{4.1}$$

on a set A, if $S^\delta((t,x,p),\Psi_k,\Psi_k(t,x,p),\Psi_{k+1})$ is positive (resp negative) for all $(t,x,p) \in A$. Let us start with the classical properties of the scheme. First, we state the monotonicity property.

Proposition 4.6. *(i) The implicit scheme is monotone in the sense of [10].*
(ii) Furthermore, for any families of functions $\Psi_k^1 \le \Psi_k^2$, $k = 0,\ldots,m$ defined on $\mathcal{D}_k$, the solutions X^1 and X^2 of the scheme at step $(t,.,p)$, starting respectively with $\Psi_k^1(t+\delta_t,.,p)$ and $\sup_{e\in E^{\delta_e}} \Psi_{k+1}^1(t,.,p\cup(t,e))$ for X^1 and $\Psi_k^2(t+\delta_t,.,p)$ and $\sup_{e\in E^{\delta_e}} \Psi_{k+1}^2(t,.,p\cup(t,e))$ for X^2 are such that:

$$X^1 \le X^2$$

where the inequality is to be taken component by component.
(iii) Finally, let δ be a given discretization. If the family of functions $\Psi_k^{1,\delta}$, $k \in \{0..m\}$ is a subsolution of the scheme and $\Psi_k^{2,\delta}$ is a supersolution of the scheme on $\mathcal{D}_k$ such that $\Psi_0^{1,\delta}(T,.) \le \Psi_0^{2,\delta}(T,.)$ on the boundary $t = T$ of the domain, then:

$$\Psi_k^{1,\delta} \le \Psi_k^{2,\delta} \text{ on } \mathcal{D}_k,\ k \in \{0..m\}$$

Proof. (i) The monotonicity in the sense of [10] is directly given by the definition of b^α and the fact that A^α is diagonally dominant for all α.
(ii) This can be proved using the fact that the matrix A^α is strictly diagonally dominant for all α. As X^1 is the solution of:

$$\min_{\alpha\in\{0,1\}^N} (A^\alpha X^1 - b^\alpha(\Psi_k^1)) = 0,$$

denoting α^1 as the minimizing quantity above, we get:

$$\begin{aligned} 0 &= A^{\alpha^1} X^1 - b^{\alpha^1}(\Psi_k^1) \\ &\le A^{\alpha^1} X^2 - b^{\alpha^1}(\Psi_k^2), \end{aligned}$$

that is:

$$A^{\alpha^1}(X^2 - X^1) \ge b^{\alpha^1}(\Psi^2) - b^{\alpha^1}(\Psi^1) \ge 0,$$

as $\Psi^2 \ge \Psi^1$ and the fact that b involves positive coefficients of Ψ. Therefore, as A^{α^1} is strictly diagonally dominant we get $X^2 \ge X^1$.
(iii) From the fact that A^α is strictly diagonally dominant, we get that at a given step $(t,.,p)$, a supersolution of (4.1) is always superior to a subsolution. Combining this result with (ii), we get that a supersolution of the scheme is superior to a subsolution on the whole domain $\mathcal{D}_k$. □

Now, we state the stability property:

Proposition 4.7. *If g, f and c are bounded by constants C_0, $C_1$and C_2, then there exists a constant C such that the solution Ψ_k^δ $k = 0, \ldots, m$ of the scheme is bounded by C for all δ, i.e. $\left\|\Psi_k^\delta\right\|_\infty \leq C$ for all δ.*

Proof. Remark that $\phi_k(t,x,p) = C_0 + C_1\left\lceil\frac{T-t_1}{h}\right\rceil + C_2(T-t)$, with $\phi_k(t,x,p) = C_0 + C_1\left\lceil\frac{T-t}{h}\right\rceil + C_2(T-t)$ is a supersolution of the scheme and of the boundary conditions for any δ. Therefore, with proposition 4.6, (ii) and (iii), we have $\Psi_k^\delta(t,x,p) \leq C_0 + C_1\left\lceil\frac{T-t_1}{h}\right\rceil + C_2(T-t)$ on $\mathcal{D}_k$, for all $k \in 0, \ldots, m$ any any step δ. We also get the lower bound by considering $-C_0 - C_1\left\lceil\frac{T-t}{h}\right\rceil - C_2(T-t)$ as a subsolution of the scheme. □

At last, we consistency is stated by the following property, in the sense on [10].

Proposition 4.8. *For all families of functions $(\phi_0, \ldots, \phi_m) \in C_b^\infty(\mathcal{D}_0) \times \ldots \times C_b^\infty(\mathcal{D}_m)$, for all $(t,x,p) \in \mathcal{D}_k$, $k = 0, ..., m$:*

$$\limsup_{\substack{\delta \to 0 \\ t',x',p' \to t,x,p \\ \xi \to 0}} \frac{S^\delta((t',x',p'), \phi_k(t',x',p'), \phi_k, \phi_{k+1})}{\delta_t} \leq F^*(t,x,p,D^2\phi_k, D\phi_k, \phi_{k+1})$$

$$\liminf_{\substack{\delta \to 0 \\ t',x',p' \to t,x,p \\ \xi \to 0}} \frac{S^\delta((t',x',p'), \phi_k(t',x',p'), \phi_k, \phi_{k+1})}{\delta_t} \geq F_*(t,x,p,D^2\phi_k, D\phi_k, \phi_{k+1})$$

Proof. This can be done considering Taylor expansions of Ψ. We will not expose the demonstration here, as it is a classical result. One can refer, for instance to [22] or [21] for such kind of impulse control problems. □

Now, we prove that the limit of the solutions of the scheme when $\delta \to 0$ converge to the value function. For any δ, let the family of functions Ψ_k^δ defined on $\mathcal{D}_k$ for $k = 0, \ldots m$, be the solution of the scheme satisfying the terminal conditions (1.3) and (1.4). We denote $\overline{\Psi_k}$ and $\underline{\Psi_k}$ the upper and lower limit of Ψ_k^δ over all sequences $\delta \to 0$:

$$\overline{\Psi_k}(t,x,p) = \limsup_{\substack{\delta \to 0 \\ (t',x',p') \to (t,x,p)}} \Psi_k^\delta(t',x',p')$$

$$\underline{\Psi_k}(t,x,p) = \liminf_{\substack{\delta \to 0 \\ (t',x',p') \to (t,x,p)}} \Psi_k^\delta(t',x',p')$$

Theorem 4.1. *Let assumptions 1.1 and 3.2 hold and $\theta = 1$. Then the solution of the implicit scheme satisfies:*
(i) For all $x \in \mathbb{R}^d$, the limit $\Psi_0(T^-, x)$ exists as:

$$\overline{\Psi_0}(T,x) = \underline{\Psi_0}(T,x) = g(x).$$

(ii) For $k = 1, \dots, m$, $p = (t_i, e_i)_{1 \le i \le k} \in \Theta_k \times E^k$, $x \in \mathbb{R}^d$:

$$\overline{\Psi_k}(t_1 + mh, x, p) = \underline{\Psi_k}(t_1 + mh, x, p) = v_{k-1}(t_1 + mh, \Gamma(x, e_1), p^-) + c(x, e_1).$$

(iii) For all $k = 0, \dots, m-1$, and $(t, x, p) \in \mathcal{D}_k^2$ such that $t = t_k + h$ or $t = T - mh$:

$$\underline{\Psi_k}(t, x, p) \ge \sup_{e \in E} \underline{\Psi_{k+1}}(t, x, p \cup (t, e)).$$

(iv) The solution of the scheme converges locally uniformly on $\mathcal{D}_k$, $k = 0, \dots, m$:

$$\overline{\Psi_k}(t, x, p) = \underline{\Psi_k}(t, x, p) = v_k(t, x, p),$$

for all $(t, x, p) \in \mathcal{D}_k$.

Proof. This proof is not meant to be exhaustive. In particular, we will not prove convergence for the discrete schemes of linear equations, and obstacle problems with an exogenous given obstacle. For the proof in these cases, one can refer to [8].
First, we prove (i). As Ψ_0^δ is the solution of the linear scheme $S^{1,\delta}$ on $(T - mh, T) \times \mathbb{R}^d$, it is well known (see [8] for instance), that the solution of this scheme converges to the solution v_0 of the equation $F^1((t, x), Dv_0, D^2 v_0) = 0$ locally uniformly on $(T - mh, T] \times \mathbb{R}^d$ as the terminal condition g is continuous. Therefore (i) holds by continuity of v_0. As a byproduct, we also obtained that (iv) holds for Ψ_0 on $(T - mh, T] \times \mathbb{R}^d$.

To prove the rest of the theorem, we use a recursion on the the sets $\mathcal{D}_k(n)$ introduced in section 2. We prove by forward induction on $n = m, \dots, N$, and by backward induction on k from $m(n) = (n - m) \wedge m$ to 0 that **(H)(n,k)** holds, where:
(H)(n,k) Statements (ii), (iii) and (iv) hold fold all $(t, x, p) \in \mathcal{D}_{k'}(n)$, for $k' = m(n), \dots, k$. We know from the previous paragraph to that **(H)(m,0)** holds. Thus the induction is initialized. Now, let us describe the induction procedure:

- Assume that **(H)(n-1,k)** holds for all $k = 0, \dots, m(n-1)$. To begin, we prove that **(H)(n,m(n))** holds. As $\mathcal{D}_{m(n)}(n) = \mathcal{D}^1_{m(n)}(n)$, we know that for each δ, the function $\Psi^\delta_{m(n)}$ is a solution of the linear scheme $S^{1,\delta}$. The terminal condition of the scheme is given by (1.3) and involves $\Psi^\delta_{m(n)-1}$ on $\mathcal{D}_{m(n)-1}(n-1)$. But by induction hypothesis **(H)(n-1,m(n-1))**, we know that $\Psi^\delta_{m(n)-1}$ converges locally uniformly to $v_{m(n)-1}$ on $\mathcal{D}_{m(n)-1}(n-1)$. That is to say, it converges uniformly on every compact. In particular, for any $0 < \eta < h$, the set:

$$\tilde{\mathcal{D}}^\eta_{m(n)-1}(n-1) = \left\{(t, x, p) \in \mathcal{D}_{m(n)-1} \ s.t. \ x \in A, \ t_1 \ge T - nh + \eta\right\}$$

is a compact set. Therefore, for any $\varepsilon > 0$, $\eta > 0$, there exists C_ε such that for all $\|\delta\|_\infty \le C_\varepsilon$

$$v_{m(n)-1} - \varepsilon \le \Psi^\delta_{m(n)-1} \le v_{m(n)-1} + \varepsilon \text{ on } \tilde{\mathcal{D}}^\eta{}_{m(n)-1}(n-1).$$

With proposition 4.7, $\Psi^\delta_{m(n)-1}$ is bounded independently of δ and (t,x,p) on $\mathcal{D}_{m(n)-1}(n-1)/\tilde{\mathcal{D}^\eta}_{m(n)-1}(n-1)$. Therefore, as we have a stable, monotone and consistent scheme for a linear equation, we get that:

$$\overline{\Psi_{m(n)}}(t_1+mh,x,p)-\varepsilon \leq v_{k-1}(t_1+mh,\Gamma(x,e_1),p^-)+c(x,e_1) \leq \underline{\Psi_{m(n)}}(t_1+mh,x,p)+\varepsilon,$$

on $\tilde{\mathcal{D}^\eta}_{m(n)}(n)$ for all η. This result is trivial on $\mathcal{D}_{m(n)}(n)/\tilde{\mathcal{D}}^0_{m(n)}(n)$ as the equation is degenerated due to assumption 3.2. Therefore, letting $\varepsilon \to 0$ and $\eta \to 0$ gives (ii) on $\mathcal{D}_{m(n)}(n)$. Then we can use the procedure in [10] together with the comparison theorem of proposition 6.3 of the former chapter and propositions 4.6, 4.7 and 4.8 to prove (iv) on $\mathcal{D}_{m(n)}(n)$. Thus, **(H)(n,m(n))** holds.

- Now, assume that **(H)(n-1,k)** hold for all $k=0,\dots,m(n-1)$, and that **(H)(n,k+1)** holds. We prove that **(H)(n,k)** holds. First, we prove that (ii) is verified on $\mathcal{D}_k(n)$. For any discretization step δ consider the solution ϕ^δ_k of the linear scheme $S^{1,\delta}$ on $\mathcal{D}_k(n)$, starting with terminal condition (3.2):

$$\phi^\delta(t,x,p) = \Psi^\delta_{k-1}(t_1+mh,\Gamma(x,e_1),p^-)+c(x,e_1),$$

for all $(x,p)\in\mathbb{R}^d\times\Theta_k(n)$, and $t\in[t_1+mh-\delta_t,t_1+mh)$. We get, as in the previous step, that the lower limit of ϕ^δ is such that:

$$\underline{\phi_k}(t_1+mh,x,p) \geq v_{k-1}(t_1+mh,\Gamma(x,e_1),p^-)+c(x,e_1).$$

Therefore, is ϕ^δ_k is a subsolution of the scheme S^δ, we get by proposition 4.6, (iii) that $\phi^\delta_k \leq \Psi^\delta_k$ on $\mathcal{D}_k$ for all k,δ, therefore:

$$\underline{\Psi_k}(t_1+mh,x,p) \geq v_{k-1}(t_1+mh,\Gamma(x,e_1),p^-)+c(x,e_1),$$

for all $(x,p)\in\mathbb{R}^d\times\Theta_k(n)$.
To prove the converse inequality, we use **(H)(n,k+1)**. We suppose that (t,x,p) is in $\mathcal{D}^2_k$, otherwise one can proceed as for the former inequality. As Ψ^δ_{k+1} converges locally uniformly to v_{k+1} on $\mathcal{D}_{k+1}(n)$ by **H(n,k+1)** , we get that for any $\varepsilon>0$,$\eta>0$ there exist C^1_ε such that for all $\|\delta\|_\infty \leq C^1_\varepsilon$:

$$\Psi^\delta_{k+1} \leq v_{m(n)-1}+\varepsilon \text{ on } \tilde{\mathcal{D}}^\eta_{k+1}(n). \tag{4.2}$$

Now, we denote, for $(t,x,p)\in\tilde{\mathcal{D}}^{2,\eta}_n(k+1)$:

$$\tilde{v_k}(t,x,p)=\sup_{e\in E} v_{k+1}(t,x,p\cup(t,e)).$$

As v_{k+1} is continuous and the maximum is taken over the compact set E, $\tilde{v_k}$ is continuous on $\tilde{\mathcal{D}}^{2,\eta}_n(k)$. Therefore, with (4.2) we get, as a classical result for obstacle problem that:

$$\overline{\Psi_k}(t_1+mh,x,p) \leq \max\Big(\quad v_{k-1}(t_1+mh,\Gamma(x,e_1),p^-)+c(x,e_1), \tag{4.3}$$
$$\tilde{v_k}((t_1^-+mh)^-,x,p)+\varepsilon\Big).$$

But with condition (1.3), we know that:

$$\tilde{v_k}((t_1^- + mh)^-, x, p) = \sup_{e \in E} v_k(t, \Gamma(x, e_1), p^- \cup (t, e)) + c(x, e_1)$$

and with the dynamic programming principle of corollary 3.1 in the former chapter we have:

$$v_{k-1}(t, \Gamma(x, e_1), p^-) \geq v_k(t, \Gamma(x, e_1), p^- \cup (t, e)),$$

for all $e \in E$. Therefore, letting $\varepsilon \to 0$ and $\eta \to 0$ in (4.3) gives the result. Thus we proved (ii) on $\tilde{\mathcal{D}}_k^{2,0}(n)$. As before, the proof on $\mathcal{D}_k^2(n)/\tilde{\mathcal{D}}_k^{2,0}(n)$ is easier as the equation is degenerated.
It remains to prove (iii). First, consider some $(t, x, p) \in \mathcal{D}_k^2(n)$ such that $t = t_k + h$. As, by an elementary property of the scheme $S^{2,\delta}$ we have $\Psi_k^\delta(t, x, p) \geq \Psi_{k+1}^\delta(t, x, p \cup (t, e))$ for all $e \in E^{\delta_e}$. Remember that, by **(H)(n,k+1)**, we have that $\Psi_{k+1}^\delta \to v_{k+1}$ uniformly on $\tilde{\mathcal{D}}_n^\eta(k+1)$, for all $\eta > 0$ and that v_{k+1} is uniformly continuous on this set. Thus, from the fact that E is a compact set, we get that for all $\varepsilon > 0$ there exists C_ε such that if $\|\delta\|_{infty} \leq C_\varepsilon$:

$$\Psi_k^\delta(t, x, p) \geq \sup_{e \in E} v_{k+1}(t, x, p \cup (t, e)) - \varepsilon$$

for all $(t, x, p) \in \tilde{\mathcal{D}}_n^{2,\eta}(k)$. Combining this fact with the properties of the linear scheme S^1 satisfied by Ψ_k on $\mathcal{D}_k^1(n)$ for $t < t_k + k$, and letting $\varepsilon \to 0$, we get that:

$$\underline{\Psi_k}(t, x, p) \geq \sup_{e \in E} v_{k+1}(t, x, p \cup (t, e)) \geq \sup_{e \in E} \overline{\Psi_{k+1}}(t, x, p \cup (t, e)).$$

for all $(t, x, p) \in \tilde{\mathcal{D}}_k^{2,\eta}(n)$ such that $t = t_k + h$. Once again, the proof on $\mathcal{D}_k^2(n)/\mathcal{D}^{\tilde{2,0}}{}_k(n)$ is simpler.
At last, the proof of (iii) for $t = T - mh$ is based on the fact that Ψ_k^δ satisfies a linear equation for $t > T - mh$, thus converges to v_k, that Ψ_{k+1}^δ converges to v_{k+1} due to hypothesis **(H)(n,k+1)**, and that the value function is itself continuous and such that:

$$v_k(T - mh, x, p) \geq \sup_{e \in E} v_k(T - mh, x, p \cup (T - mh, e))$$

on all $(T - mh, x, p) \in \mathcal{D}_k^2$.
Finally, to prove the convergence of the algorithm on $\mathcal{D}_k(n)$, we use the method of [10]. Then convergence follows from the monotonicity, stability and consistence properties of propositions 4.6, 4.7 and 4.8. To apply the comparison principle of proposition 6.3 in the former chapter, we use the fact that properties (ii) and (iii) satisfied on $\mathcal{D}_{\hat{k}}(n-1)$ for all $\hat{k} = 0, \ldots, m(n-1)$ and on $\mathcal{D}_{\hat{k}}(n)$ for all $\hat{k} = k, \ldots, m(n)$. Then (iv) is proved on $\mathcal{D}_k(n)$, and **(H)(n,k)** is satisfied.

□

5 Impulse control problem without delay

The major practical issue of the numerical procedure is the dimension of the problem. Indeed, the dimension of the state space $\mathcal{D}_m$ is $1 + d + m(1 + dim(E))$. This dimension grows linearly with m, which is therefore a critical parameter. Thus, in the practical problem considered on the following, we will only deal with $m = 1$ in order to obtain reasonable computational time. On the other hand, we try to estimate the consequences of execution delay only, and not the consequences of discrete control. As the minimum time between to order can be written as $h = \frac{d}{m}$, the discrete behavior of the control cannot be neglected with respect to the delay.
To circumvent this difficulty, and to avoid taking high values of m, we also consider the same optimization problem with discrete control but without execution delay. In this problem, the process X follows the same diffusion, but this time, the actions of the agent, decided at any stopping time τ_i take effect immediately:

$$X_{\tau_i} = \Gamma(X_{\tau_i^-}, \xi_i).$$

Therefore, for a given control $\alpha \in \mathcal{A}$, the controlled process X^α is defined as the solution of the s.d.e:

$$X_s^\alpha = X_0 + \int_0^s b(X_u^\alpha)du + \int_0^s \sigma(X_u^\alpha)dW_u + \sum_{\tau_i \le s} \big(\Gamma(X^\alpha_{(\tau_i^-}, \xi_i) - X^\alpha_{(\tau_i)^-}\big).$$

the set of admissible controls is written as:

$$\mathcal{A} = \Big\{\alpha = (\tau_i, \xi_i)_{i\ge 1} : \tau_i \text{ is a s.t., } \xi_i \text{ is } \mathcal{F}_{\tau_i}\text{adapted}, \tau_{i+1} - \tau_i \ge h\Big\}.$$

And the objective is to maximize the expectation:

$$\mathbb{E}\Big[\int_0^T f(X_t)dt + g(X_T) + \sum_{\tau_i \le T} c(X_{\tau_i^-}, \xi_i)\Big].$$

This kind of problem has been studied in [13] and more recently in [56]. The reader can refer to these works for some ideas on the proofs of the following results. Here, we just state some results without demonstration.
We define u as the value function of the problem. It can be written as a function of time t, of the variable $x = X_t$, and of the time t_1 when the last order was passed. This function is defined on the following set:

$$\tilde{\mathcal{D}} = \Big\{(t, x, t_1) \in [0, T] \times [0, +\infty)^d \times [0, T] \mid t_1 \le t\Big\}$$

We divide it into two subsets:

$$\tilde{\mathcal{D}}^1 = \Big\{(t, x, t_1) \in \tilde{\mathcal{D}} \mid t - t_1 < h\Big\}$$
$$\tilde{\mathcal{D}}^2 = \Big\{(t, x, t_1) \in \tilde{\mathcal{D}} \mid t - t_1 \ge h\Big\}$$

In this case, classical dynamic programming argument leads to the proof that u is a viscosity solution of the equation:

$$-\frac{\partial u}{\partial t}(t,x,t_1) - \mathcal{L}u(t,x,t_1) - f(x) = 0 \tag{5.1}$$

on $\tilde{\mathcal{D}}^1$, and of the equation:

$$\min\Big\{ \qquad -\frac{\partial u}{\partial t}(t,x,t_1) - \mathcal{L}u(t,x,t_1) - f(x) \tag{5.2}$$

$$, u(t,x,t_1) - \sup_{e\in E}\left\{u(t,\Gamma(x,e),t) + c(x,e)\right\}\Big\} \;=\; 0 \tag{5.3}$$

on $\tilde{\mathcal{D}}^2$.

In numerous cases the terminal condition for the value function is of the form:

$$u(T^-,x,t_1) = 1_{t_1>T-h}g(x) + 1_{t_1\le T-h}\max\Big\{g(x), \sup_{e\in E}\left\{g(\Gamma(x,e)) + c(x,e)\right\}\Big\}$$

But here, with assumption 2.7 in the former chapter, we can prove easily that its is not optimal to pass an impulsion at time T, therefore the boundary condition reduces to:

$$u(T^-,x,t_1) = g(x) \tag{5.4}$$

6 Optimal investment and indifference pricing problems

6.1 Problem formulation

Here, we restate the financial example of the former chapter, and we slightly modify it for our numerical purpose. In the following, we will consider the delayed impulse control problem, but the problem of discrete hedging without delay of section 5 can be recovered by taking $m = 0$ in this description. We consider a two-asset, one factor market model consisting a cash account and a risky asset. We take the cash account as a numeraire, and we assume that the price of the risky asset follows a Black Scholes model:

$$dS_t = S_t\left(\mu dt + \sigma dW_t\right)$$

We denote by Y_t the number of shares in the stock, and by Z_t the amount of money (cash holdings) held by the investor at time t. We assume that the investor can only trade discretely, and his orders are executed with delay. This is modeled through an impulse control $\alpha = (\tau_i,\xi_i)_{i\ge1} \in \mathcal{A}$, where τ_i are the decision times, and ξ_i are the numbers of stock the agent decides to possess at τ_i, but that he will have at times $\tau_i + mh$. Therefore, the agent will buy $\xi_i - Y_{\tau_i+mh}$ shares at time $\tau_i + mh$ if $\xi_i > Y_{\tau_i+mh}$ and will sell $Y_{\tau_i+mh} - \xi_i$ shares if $Y_{\tau_i+mh} < \xi_i$. The dynamics of Y are then given by

$$Y_t \;=\; Y_0 + \sum_{\tau_i+mh\le t}(\xi_i - Y_{(\tau_i+mh)^-}),$$

which means that discrete trading $\Delta Y_t := \xi_i - Y_{t^-}$ occur at times $s = \tau_i + mh$, $i \ge 1$. We assume that there are fixed minimal and a maximal number of shares that the agent can hold, that is:

$$Y \in [y_{min}, y_{max}] \text{ and } \xi_i \in [y_{min}, y_{max}] \tag{6.1}$$

In absence of trading, the discounted cash holdings Z is constant. When a discrete trading ΔY_t occurs, this results in a variation of cash holdings by $\Delta Z_t := Z_t - Z_{t^-} = -(\Delta Y_t)S_t$, from the self-financing condition. In other words, the dynamics of Z are given by

$$Z_t = Z_0 - \sum_{\tau_i+mh\leq t} (\xi_i - Y_{(\tau_i+mh)^-}).S_{\tau_i+mh}.$$

The wealth process is equal to $L(S_t, Y_t, Z_t) = Z_t + Y_t S_t$. This financial example corresponds to the general model with $X = (S, Y, Z)$, $b = (\beta\ 0\ r)'$, $\sigma = (\gamma\ 0\ 0)$, and

$$\Gamma(s, y, z, e) = \begin{pmatrix} s \\ e \\ z + (y - e)s \end{pmatrix}.$$

Note that one could introduce fixed and proportional transaction costs by modifying Γ, and that assumption 2.7 of the former chapter would still be satisfied. It remains to fix the objective of the financial agent, it is given by the utility of the liquidative value of his portfolio at time T. In the case of indifference pricing, the agent has sold a number κ of options which pay $H(S_T)$ at maturity. Therefore, the objective of the agent is to maximize:

$$V_0(z, s, y, \kappa) = \sup_{\alpha\in\mathcal{A}} \mathbb{E}\Big[U\big(Z_T + Y_T S_T - \kappa H(S_T)\big)\Big]$$

If $\kappa = 0$, this is an optimal investment problem. The utility indifference ask price $\pi_a(\kappa, z)$ is the price at which the investor is indifferent (in the sense that her expected utility is unchanged under optimal trading) between paying nothing and not having the claim, and receiving $\pi_a(\kappa, z)$ now to deliver κ units of claim at time T. It is then defined as the solution of

$$V_0(z + \pi_a(\kappa, z), \kappa) = V_0(z, 0).$$

6.2 Considering exponential utility

In the numerical computations we performed, we chose the exponential utility:

$$U(x) = -e^{-\gamma x}.$$

This utility has the nice property that $U(z + z') = e^{-\gamma z'}U(z)$. This immediately leads to:

$$V_0(z, s, y, \kappa) = e^{-\gamma z}V_0(0, s, y, \kappa). \tag{6.2}$$

The drawback is that this utility function does not follow the linear growth property for large losses. Nevertheless, one can circumvent this difficulty by considering a bounded Black Scholes model. We make the following assumption:

Assumption 6.4. *The spot price S is stopped as soon as it reaches a given barrier s_{max}.*

This can be interpreted as the fulfillment of assumption 3.2. But, as pointed out by [8] this corresponds also, in the numerical scheme, to a Dirichlet condition at the boundary $s = s_{max}$. This condition can be derived as:

$$v(t, z, s_{max}, y, p) = -exp\left(z + ys_{max} - \kappa H(s_{max})\right).$$

Furthermore, the portfolio constraint (6.1) involves that, due to the discrete hedging, the liquidative value of the agent portfolio is bounded from below by:

$$Z_T + Y_T S_T \geq z - (y_{max} - y_{min}) s_{max} \frac{T}{h}. \tag{6.3}$$

The fact that the payoff H of the option is continuous, it admits a maximum M_H on $[0, s_{max}]$ Thus, one could consider the terminal condition:

$$\tilde{U}(x) = -exp\left(-\gamma(x \vee (-\kappa M_H - (y_{max} - y_{min}) s_{max} \frac{T}{h})\right).$$

With relation (6.3), we get that the value function of the problem with terminal condition U and $\tilde{U}$ are the same, along the path (Z_t, S_t, Y_t) of the portfolio of the agent, starting from $z = 0$ at time $t = 0$. Thus, we can still use relation (6.2). Meanwhile, $\tilde{U}$ is bounded, which means that we are in the framework of the former chapter. We obtain the following expression for boundary condition:

$$v_k((t_1 + mh)^-, (s, y, 0), p) \quad = \quad e^{-\gamma((y-e_1)s)} v_{k-1}(t_1 + mh, (s, e_1, 0), p_-)$$

Therefore, we obtain a framework in which we can omit the variable z in the numerical discretization, considering only the case $z = 0$.

6.3 Numerical results

Here, we implemented the algorithm in the case $m = 1$. We chose to study two different problems: the optimal investment problem and the indifference price for a call option. We will see that the results are strongly dependent of the initial condition, in particular to the initial number of shares Y_0 held by the agent at time $t = 0$. As we will use the exponential utility described above, we can consider that the agent starts with initial wealth $Z_0 = 0$. We have to compute the value of the following functions:

$$v_0(t, (s, y)) \text{ and } v_1(t, (s, y), (t_1, e_1))$$

for the delay control problem, and the function:

$$u(t, (s, y), t_1)$$

for the impulse control problem without delay.

Optimal investment problem

Here, we compare the delay controlled problem with two other ones: the classical Merton problem, and the impulse control problem of section 5. We suppose the following set of parameters:

$$\mu = 6\%, \; \sigma = 10\%, \; s_{max} = 2, \; s_0 = 1, \; \gamma = 20, \; y_{min} = 0, \; y_{max} = 1.$$

We have $m = 1$, and we choose a delay and a minimum lag between to intervention of two month, that is $h = \frac{1}{6}$. We use an explicit scheme (that is $\theta = 0$). The space discretization

for s is done with a step $\delta_s = 0.015$, the time step is 0.001. Changing the maturity T of the problem, we obtain the values given in table 5.1 and plotted on figure 5.1 for v_0 and u with $(t,(s,y)) = (0,(1,0))$.
Note that these numerical computations are made in reasonable time, that is approximately 30 seconds for a one year maturity.
To understand the behavior of the value function better, we now consider the supremum of the value functions v_0 and u over all possible initial number of shares y. Of course, we compensate the price of these shares with the initial wealth z. That is, we calculate, for the delay control problem

$$\sup_{y\in[y_{min},y_{max}]} \left\{exp(\gamma y s_0)v_0(0,(s_0,y))\right\},$$

and for the non delayed case:

$$\sup_{y\in[y_{min},y_{max}]} \left\{exp(\gamma y s_0)u(0,(s_0,y),-h)\right\}.$$

This will help to separate the loss utility due to the delay during the problem from the loss of utility due to suboptimal initial portfolio (that is the loss of utility due to the non exposure to the risky asset during the two first month in the delayed case). The loss of utility due to the discrete or delayed investment is plotted on figure 5.2 for various maturities. These results are given in table 5.2.

We see that most of the loss of utility was mainly due to suboptimal initial conditions. This is not surprising, knowing that the optimal strategy in the Merton case is to maintain a constant amount of money invested in the risky asset. This is not very far from the case when the agent does not pass any order, which implies a constant investment in the asset in terms of number of shares.

Indifference pricing problem

Now, we consider the problem of indifference pricing. In this problem, we use the following set of parameters:

$$\mu = 0,\ \sigma = 10\%,\ s_{max} = 2,\ s_0 = 1,\ \gamma = 20,\ y_{min} = 0,\ y_{max} = 1.$$

We consider a call option of strike $K = 1$ and we compute the indifference selling price of one unit of this option, that is $\kappa = 1$. Notice that we took $\mu = 0$, so that the investment problem gets degenerated for $\kappa = 0$, that is $v_0(0,(s_0,0)) = -1$. We use the same discretization as before. First, we compute the indifference price of a 3 years call option, for $y_0 = 0$, for various delays. We obtain the prices of table 5.3, plotted on of figure 5.3.

Now, as in the previous example, we compute the value function with the best initial endowment y. This is important, as in practice, a bank would sell an at the money forward starting call. This means that, after the option is sold, the strike of the call option would be determined later at some striking date, the strike depending of the price of the underlying. This is in order for the bank to have purchased the correct amount of shares of underlying at the striking date. We the results on figure 5.4. We see that a large part of the difference between the Black Scholes price and the price with delay has disappeared, but that the

maturity	Continuous investment	discrete investment	delayed investment	relative loss of utility due discrete invest	relative loss of utility due to delay
0.3	-0.951	-0.952	-0.977	0.01%	2.74%
0.5	-0.920	-0.920	-0.944	0.02%	2.68%
0.75	-0.883	-0.883	-0.906	0.04%	2.73%
1	-0.847	-0.847	-0.869	0.05%	2.70%
1.5	-0.779	-0.779	-0.800	0.07%	2,71%
2	-0.717	-0.717	-0.737	0.10%	2,83%
2.5	-0.659	-0.660	-0.677	0.12%	2,73%
3	-0.607	-0.608	-0.624	0.15%	2,87%
4	-0.514	-0.515	-0.529	0.20%	2,88%
5	-0.435	-0.436	-0.448	0.25%	2,90%

Table 5.1: delayed and discrete investment problem with no initial endowment

maturity	Continuous investment	discrete investment	delayed investment	relative loss of utility due discrete invest	relative loss of utility due to delay
0.3	-0.951	-0.951	-0.952	0.002%	-0.018%
0.5	-0.920	-0.920	-0.920	0.004%	-0.032%
0.75	-0.883	-0.883	-0.883	0.006%	-0.049%
1	-0.847	-0.847	-0.847	0.008%	-0.066%
1.5	-0.779	-0.779	-0.780	0.012%	-0.102%
2	-0.717	-0.717	-0.718	0.017%	-0.137%
2.5	-0.659	-0.660	-0.661	0.021%	-0.173%
3	-0.607	-0.607	-0.608	0.026%	-0.208%
4	-0.514	-0.514	-0.515	0.034%	-0.280%
5	-0.435	-0.436	-0.437	0.042%	-0.349%

Table 5.2: delayed and discrete investment problem with optimal initial endowment

effect of delay is still non negligible, contrarily to the previous example.
Finally, we perform the same calculations, taking a constant delay of 2 months, and considering various maturities, with the best initial endowment, it leads to the results of figure 6.3, and table 5.4.

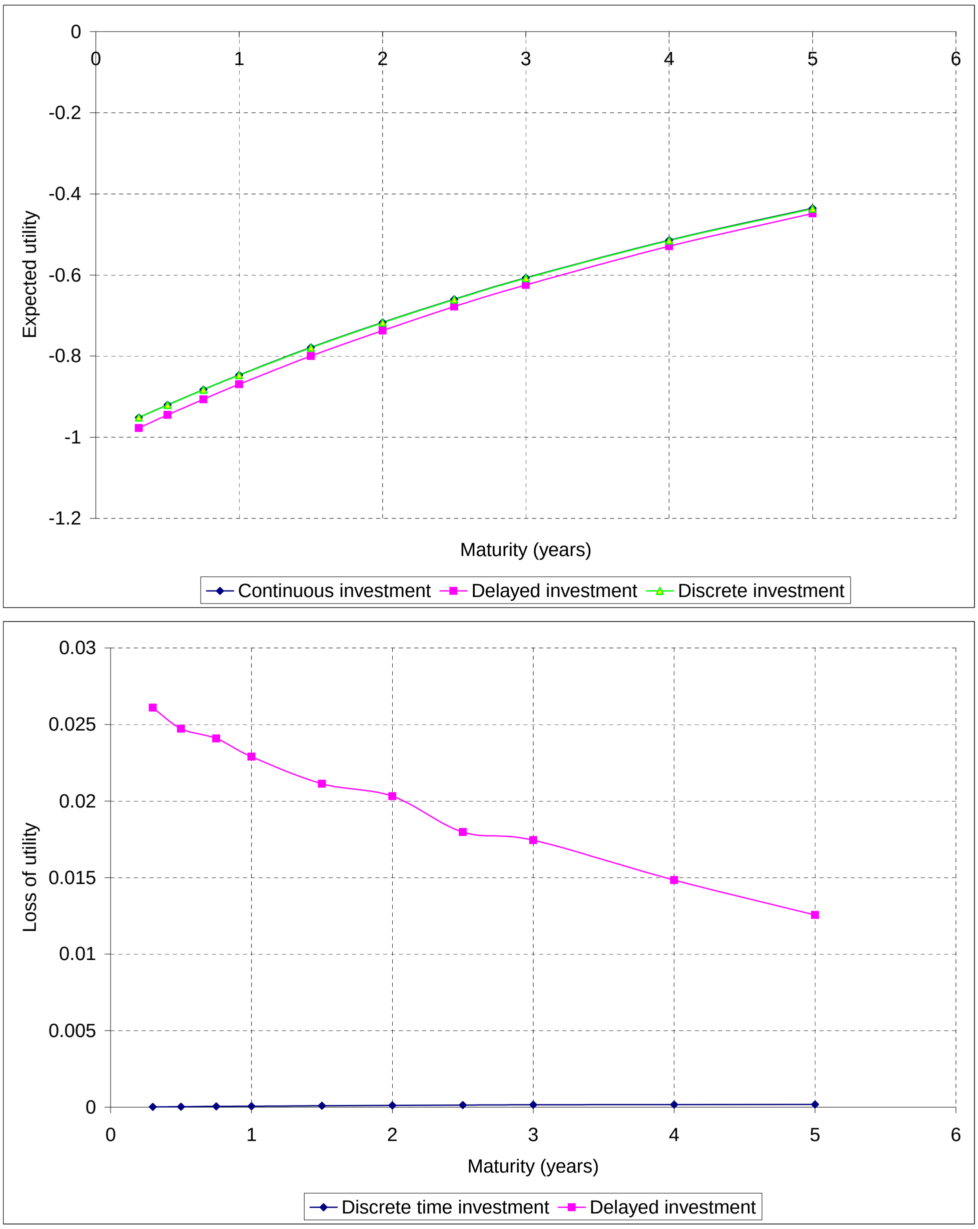


Figure 5.1: Expected utility for the Merton problem, the Discrete problem and the Delay problem with no initial endowment (above), and difference w.r.t. the Merton case (below)

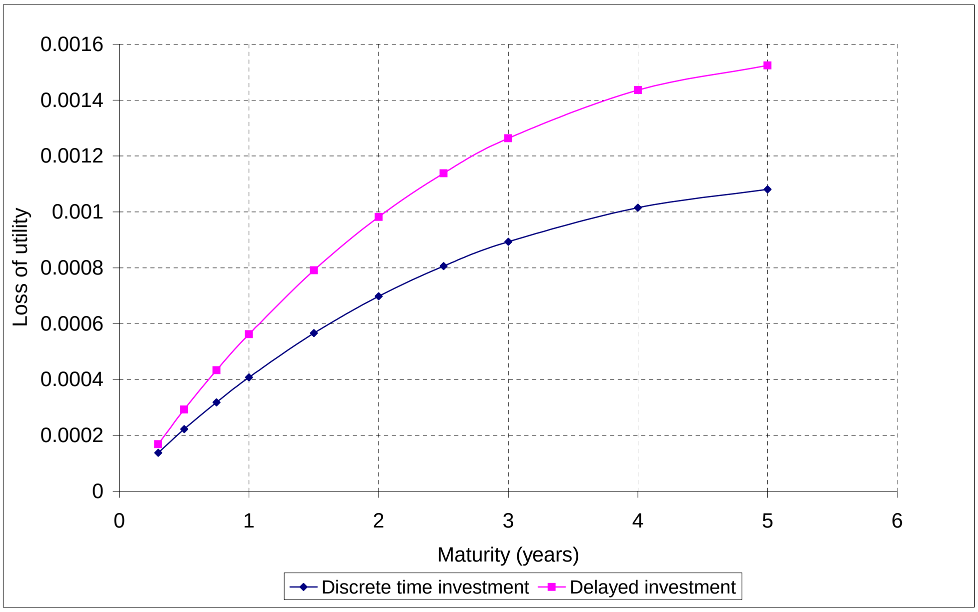


Figure 5.2: Difference w.r.t. the Merton case (below) for discrete and delayed investment problem with optimal initial endowment in risky asset.

delay (years)	BS price	discrete hedging $y = 0$	delayed hedging $y = 0$	discrete hedging optimal y	delayed hedging optimal y
0.01	6.90%	6.94%	6.94%	6.85%	6.86%
0.025	6.90%	6.94%	7.03%	6.87%	6.91%
0.05	6.90%	6.97%	7.19%	6.89%	6.97%
0.075	6.90%	6.99%	7.34%	6.92%	7.05%
0.1	6.90%	7.03%	7.48%	6.94%	7.11%
0.15	6.90%	7.08%	7.79%	6.98%	7.23%
0.2	6.90%	7.16%	8.16%	7.03%	7.35%
0.3	6.90%	7.26%	8.75%	7.11%	7.59%
0.4	6.90%	7.42%	9.58%	7.19%	7.81%
0.5	6.90%	7.53%	10.32%	7.27%	8.02%
0.6	6.90%	7.66%	10.98%	7.35%	8.22%
0.7	6.90%	7.80%	11.84%	7.42%	8.41%
0.8	6.90%	7.93%	12.86%	7.49%	8.58%
0.9	6.90%	8.12%	13.97%	7.56%	8.75%
1	6.90%	8.48%	15.60%	7.62%	8.90%
1.5	6.90%	8.97%	23.49%	7.89%	9.47%

Table 5.3: indifference price for different values of h, in percentage of the initial spot price

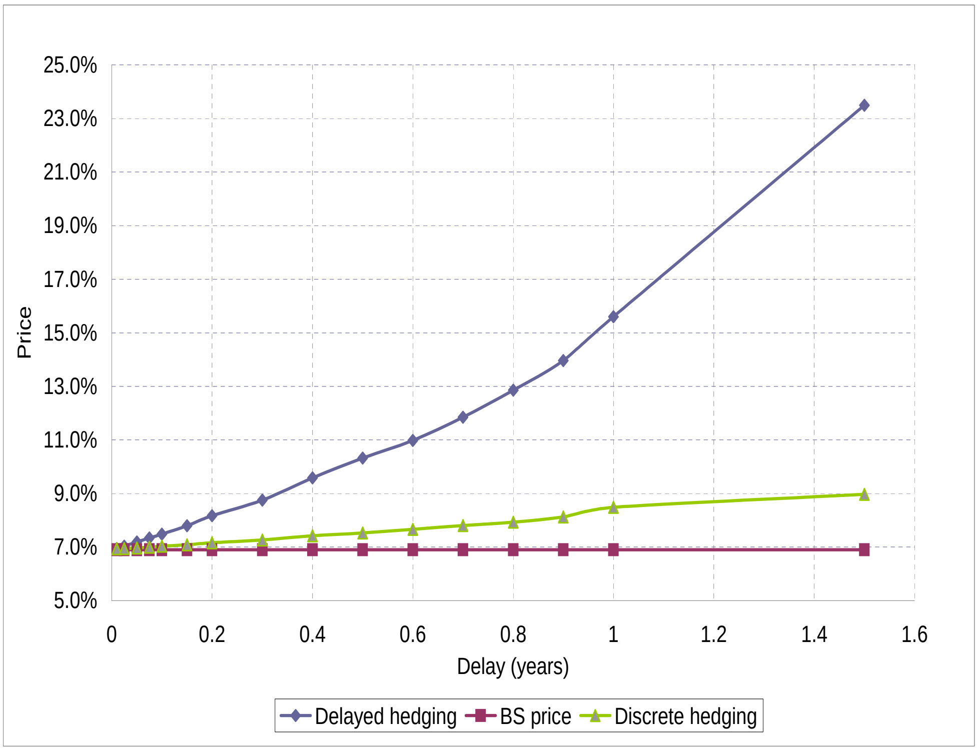


Figure 5.3: Indifference price for discrete and delayed hedging for various h and a 3 years ATM call option, with no initial endowment.

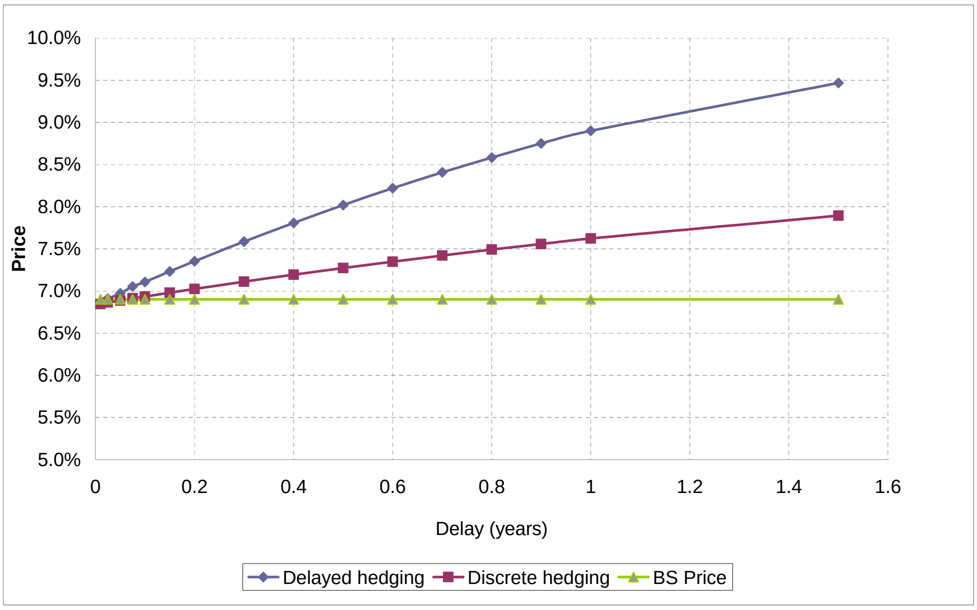


Figure 5.4: Indifference price for discrete and delayed hedging for a 3 years ATM call option, with optimal initial endowment in risky asset.

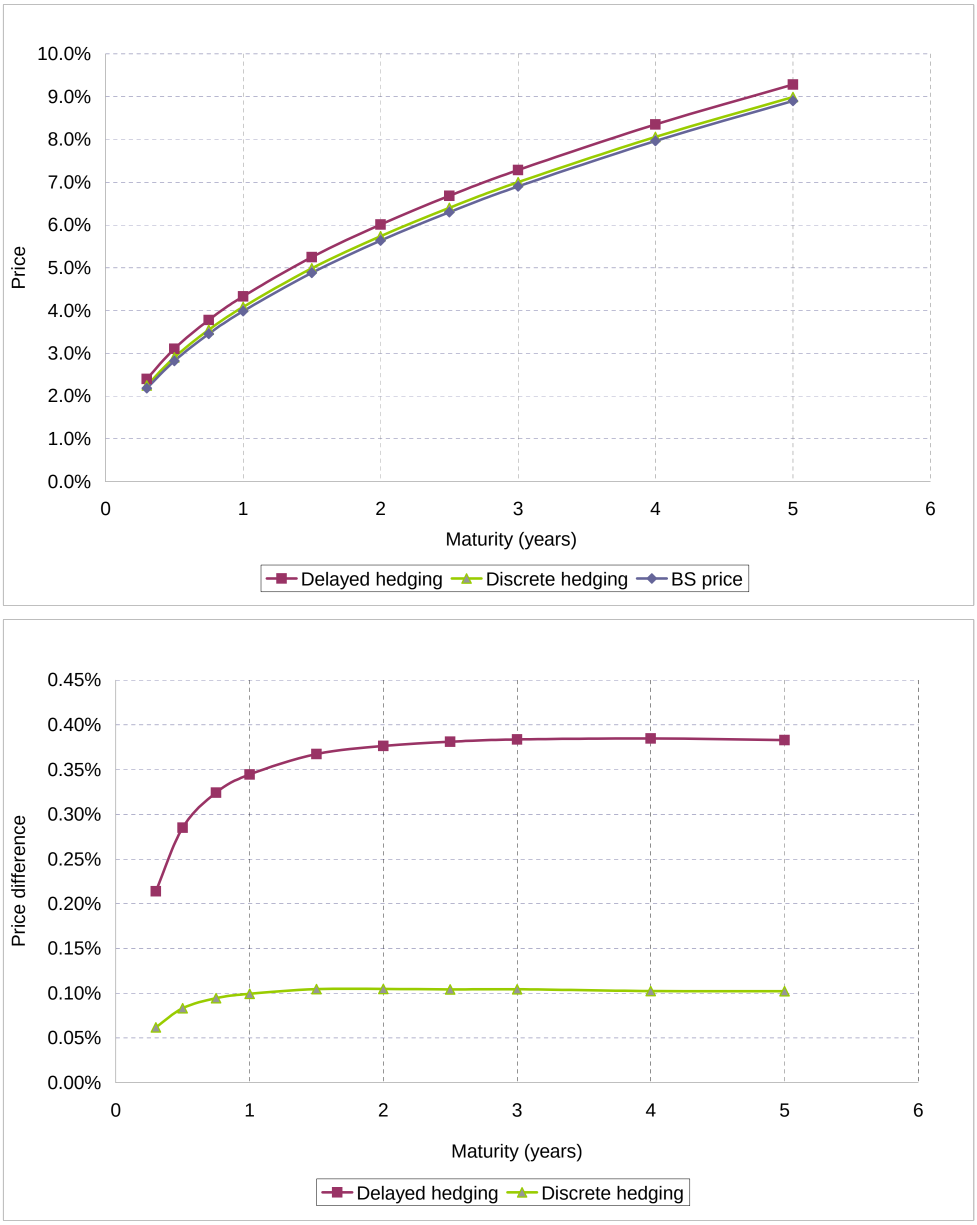


Figure 5.5: Indifference price for discrete and delayed hedging for h=2 month, with optimal endowment in risky asset (above) and difference w.r.t. the BS price (below).

maturite (years)	BS price	discrete hedging	difference w.r.t. BS	delayed hedging	difference w.r.t. BS
0.3	2.18%	2.25%	0.06%	2.40%	0.21%
0.5	2.82%	2.90%	0.08%	3.11%	0.29%
0.75	3.45%	3.55%	0.09%	3.78%	0.32%
1	3.99%	4.09%	0.10%	4.33%	0.34%
1.5	4.88%	4.99%	0.10%	5.25%	0.37%
2	5.64%	5.74%	0.10%	6.01%	0.38%
2.5	6.30%	6.40%	0.10%	6.68%	0.38%
3	6.90%	7.00%	0.10%	7.28%	0.38%
4	7.97%	8.06%	0.10%	8.35%	0.38%
5	8.90%	8.99%	0.10%	9.29%	0.38%

Table 5.4: indifference price for different maturities, constant $h = 2$ month, with optimal initial endowment in risky assets.

Bibliography

[1] M. Akian, J.L. Menaldi, and A. Sulem. On an investement-consumption model with transaction costs. *SIAM J. Control Optim.*, 34:329–364, 1996.

[2] L. Alvarez and J. Keppo. The impact of delivery lags on irreversible investment under uncertainty. *European Journal on Operational Research*, 136:173–180, 2002.

[3] M. Avellaneda, A. Levy, and A. Paras. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2, 1995.

[4] A. Bar-Ilan and W. Strange. Investment lags. *American Economic Review*, 86:610–622, 1996.

[5] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems and Control: Foundations and Applications. Birkhäuser, Boston, 1997.

[6] G. Barles. An approach of deterministic control problems with unbounded data. *Ann. Inst. H. Poincaré, Anal. Non Linéaire*, 7(4):235–258, 1990.

[7] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*, volume 17 of *Mathématiques et Applications*. Springer, Paris, 1994.

[8] G. Barles, Ch. Daher, and M. Romano. Convergence of numerical schemes for parabolic equations arising in finance theory. *Math. Models Methods Appl. Sci.*, 5(1):125–143, 1995.

[9] G. Barles and E.R. Jakobsen. On the convergence rate of approximation schemes for hamilton-jacobi-bellman equations. *Mathematical Modelling and Numerical Analysis*, 36(1):33–54, 2002.

[10] G. Barles and P.E. Souganidis. Convergence of approximation schemes for fully non-linear second order equations. *Asymptotic Analysis*, 4:271–283, 1991.

[11] E. Bayraktar and M. Egami. The effects of implementation delay on decision-making under uncertainty. *Stochastic Processes and their Applications*, To appear.

[12] I. Ben Tahar and B. Bouchard. Explicit characterization of the super-replication strategy in financial markets with partial transaction costs. *Stochastic Processes and their Applications*, 117(5):655–672, 2007.

[13] A. Bensoussan and J.L. Lions. *Application des inéquations variationnelles en contrôle stochastique*, volume 6 of *Méthodes mathématiques de l'informatique.* Dunod, Paris, 1978.

[14] A. Bensoussan and J.L. Lions. *Contrôle impulsionnel et inéquations variationnelles.* Dunod, 1982.

[15] D. Bertsekas and S. Shreve. *Stochastic optimal control : the discrete time case.* Math. in Sci. and Eng., Academic Press, 1978.

[16] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[17] J.F. Bonnans, E. Ottenwaelter, and H. Zidani. Numerical schemes for the two dimensional second-order hjb equation. *ENSAIM: M2AN*, 38:723–735, 2004.

[18] J.F. Bonnans and H. Zidani. Characterization of consistency of some numerical schemes for the stochastic hjb equation. In J.L. Menaldi E. Rofman and A. Sulem, editors, *Optimal Control and PDE - Innovations et Applications.* IOS Press, 2000.

[19] B. Bouchard and N. Touzi. Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications*, 111(2):175–206, 2004.

[20] B. Bruder. Super-replication of european options with a derivative asset under constrained finite variation strategies.

[21] J.P. Chancelier, B. Øksendal, and A. Sulem. Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control. *Tr. Mat. Inst. Steklova*, 237(Stokhast. Finans. Mat.):149–172, 2002.

[22] Z. Chen and P.A. Forsyth. A numerical scheme for the impulse control formulation for pricing variable annuities with a guaranteed minimum withdrawal benefit (gmwb). 2007.

[23] P. Cheridito, H.M. Soner, and N. Touzi. The multi-dimensional super-replication problem under gamma constraints. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 22(5):633–666, 2005.

[24] P. Cheridito, H.M. Soner, and N. Touzi. Small time path behavior of double stochastic integrals and applications to stochastic control. *Ann. Appl. Probab.*, 15(4):2472–2495, 2005.

[25] R. Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16:519–542, 2006.

[26] R. Cont and J. Da Fonseca. Dynamics of implied volatility surfaces. *Quantitative Finance*, 2(1):45–60, 2002.

[27] M. Crandall, H. Ishii, and P.L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27:1–67, 2000.

[28] J. Cvitanic, H. Pham, and N. Touzi. A closed-form solution for the problem of super-replication under transaction costs. *Finance and Stochastics*, 3:35–54, 1999.

[29] J. Cvitanic, H. Pham, and N. Touzi. Super-replication in stochastic volatility models with portfolio constraints. *Journal of Applied Probability*, 36:523–545, 1999.

[30] R. Dautray and J.L. Lions. *Mathematical analysis and numerical methods for science and technology. Volume 5 . Evolution Problems I.* Springer, 2000.

[31] M. Davis, V.G. Panas, and T. Zariphopoulou. European option pricing with transaction costs. *SIAM J. Control and Optimization*, 31:470–493, 1993.

[32] L. Denis and C. Martini. A theorical framework for the pricing of contingent claims in the presence of model uncertainty. *prépublication de l'Université du Maine*, 2002.

[33] P. Dupuis and H.J. Kushner. *Numerical Methods for Stochastic Control Problems in Continuous Time.* Springer-Verlag, 1992.

[34] N. El Karoui. *Les aspects probabilistes du contrôle stochastique.* Springer-Verlag, 1981.

[35] N. El Karoui, M. Jeanblanc, and S.E. Shreve. Robustness of the black and scholes formula. *Mathematical finance*, 8:93–126, 1998.

[36] N. El Karoui and M.C. Quenez. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control Optim.*, 33(1):29–66, 1995.

[37] A. Erdélyi, M.F. Oberhettinger, and F.G. Tricomi. *Tables of Integral Transforms. Based, in Part, on Notes Left by Harry Bateman and Compiled by the Staff of the Bateman Manuscript Project*, volume 2. McGraw-Hill, 1954.

[38] W. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, 1993.

[39] H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. *Applied Stochastic Analysis*, pages 389–414, 1990.

[40] F. Gozzi and T. Vargiolu. On the superreplication approach for european multiasset derivatives. *Mathematical methods of operations rasearch*, 1(55):69–91, 2002.

[41] R.L. Graham, D.E. Knuth, and O. Patashnik. *Concrete Mathematics, A Foundation For Computer Science*. Addison-Wesley, Paris, 1994.

[42] K. Helmes. The "local" law of the iterated logarithm for processes related to lévy's stochastic area process. *Studia Mathematica*, 84:229–237, 1986.

[43] S. Heston. A closed-form solution for option with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6, 1993.

[44] J.E. Hodder, A. Tourin, and T. Zariphopoulou. Numerical schemes for variational inequalities arising in international asset pricing. *Computational Economics*, 17(1):43–80, 2001.

[45] S.D. Hodges and A. Neuberger. Optimal replication of contingent claims under transaction costs. *Review of Futures Markets*, 8:222–239, 1989.

[46] R.A. Howard. *Dynamic Programming and Markov Process*. MIT, 1960.

[47] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 3, 1987.

[48] H. Ishii and P.L. Lions. Viscosity solutions of fully non-linear second-order partial differential equations. *J. Differential Equations*, 83:26–78, 1990.

[49] B. Lapeyre, A. Sulem, and D. Talay. *Simulation of Financial Models: Mathematical Foundations and Applications.* Cambridge University Press, to appear.

[50] P. Levy. Wiener's random function, and other laplacian random functions. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 171–187, 1950.

[51] P.L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 2: viscosity solutions and uniqueness. *Communications in partial differential equations*, 8:1229–1276, 1983.

[52] V. Ly Vath, M. Mnif, and H. Pham. A model of portfolio selection under liquidity risk and price impact. *Finance and Stochastics*, 11:51–90, 2007.

[53] R. Merton. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory*, 3:373–413, 1971.

[54] M. Musiela and T. Zariphopoulou. An example of indifference prices under exponential preferences. *Finance Stoch.*, 8(2):229–239, 2004.

[55] B. Oksendal and A. Sulem. Optimal stochastic impulse control with delayed reaction. Preprint, University of Oslo.

[56] B. Oksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions.* Universitext. Springer Verlag, 2005.

[57] H. Pham. On quadratic hedging in continuous time. *Mathematical Methods of Operations Research*, 51:315–339, 2000.

[58] H. Pham. On some recent aspects of stochastic control and their applications. *Probability Surveys*, 2:1–549, 2005.

[59] H. Pham. *Optimisation et controle stochastique appliques a la finance*. Springer-Verlag, To Appear.

[60] M. Robin. Contrôle impulsionnel avec retard pour les processus de diffusion. *C.R.A.S*, 282:463–466, 1976.

[61] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10:259–276, 2000.

[62] M. Schweizer. A guided tour through quadratic hedging approaches. In M. Musiela E. Jouini, J. Cvitanic, editor, *Option Pricing, Interest Rates and Risk Management*. Cambridge University Press, 2001.

[63] H.M. Soner and N. Touzi. Super replication under gamma constraints. *SIAM Journal on Control and Optimization*, 39(1):73–96, 2000.

[64] H.M. Soner and N. Touzi. Stochastic target problems, dynamic programming and viscosity solutions. *SIAM Journal on Control and Optimization*, 41:404–424, 2002.

[65] H.M. Soner and N. Touzi. Hedging under gamma constraints by optimal stopping and face-lifting. *Math. Finance*, 17(1):59–79, 2007.

[66] A. Subramanian and R. Jarrow. The liquidity discount. *Mathematical Finance*, 11:447–474, 2001.

[67] A. Tourin and T. Zariphopoulou. Numerical schemes for investment models with singular transactions. *Computational Economics*, 7(4):287–307, 1994.

Résumé : Nous étudions quelques applications du contrôle stochastique à la couverture d'options en présence d'illiquidité. Dans la première partie, nous nous intéressons à un problème de surcouverture d'option dans un modèle à volatilité stochastique. L'originalité provient du fait que l'actif servant à couvrir la volatilité n'est pas liquide et que l'agent devra donc opérer un montant total fini de transactions. La deuxième partie concerne la couverture d'option en présence de volatilité incertaine dont la dynamique n'est pas spécifiée. Nous introduisons un critère permettant d'obtenir des prix d'options non triviaux, en autorisant l'agent à perdre de l'argent pour des réalisations de la volatilité qu'il juge peu probables. Enfin dans une troisième partie nous étudions un problème de contrôle impulsionnel pour lequel les contrôles prennent effet avec retard. Cette étude s'applique notamment à la couverture d'options sur hedge funds, pour lesquels les ordres d'achat et de vente sont exécutés avec retard. Dans chaque partie, nous caractérisons la fonction valeur du problème comme étant l'unique solution de viscosité d'une équation aux dérivées partielles. Dans la première et la troisième partie, nous introduisons dans un second chapitre des algorithmes de résolution numériques de ces EDP par différences finies. La convergence de ces algorithmes est prouvée de manière théorique.

Mots-clés : contraintes gamma, surréplication, solutions de viscosité, intégrales stochastiques doubles, volatilité incertaine, contrôle impulsionnel, retard d'execution, principe de comparaison, différences finies.

Discipline : Mathématiques

Abstract : We study some applications of stochastic control to option hedge with illiquidity. In the first part, we focus on a superreplication problem in a stochastic volatility model. The specificity comes from the fact that the asset which is used to hedge volatility is illiquid, thus only a finite total amount of transactions can be operated during the hedging. The second part is about option hedging in presence of uncertain volatility, which dynamics are unspecified. We introduce a criterion to obtain non trivial prices, by allowing the agent to lose money for improbable volatility scenarios. At last, in the third part, we study an impulse control problem in which the actions take effect with delay. This can be applied for hedging options on hedge funds. Indeed, buying and selling orders on these funds are executed with delay. In each part, we characterize the value function of the problem as the unique viscosity solution of a partial differential equation. In the first and third parts, we also introduce, in a second chapter, numerical algorithms to solve those PDE with finite differences methods. Convergence of these algorithms is proved in a theoretical framework.

Key words : gamma constrains, super replication, viscosity solutions, double stochastic integral, uncertain volatility, impulse control, execution delay, comparison principle, finite differences.

Laboratoire de Probabilités et Modèles Aléatoires,
CNRS-UMR 7599, UFR de Mathématiques, case 7012
Université Paris 7, Denis Diderot
2, place Jussieu, 75251 Paris Cedex 05.